

Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation

Marcus T. Pearce^{a,b}, María Herrojo Ruiz^c, Selina Kapasi^d, Geraint A. Wiggins^a, Joydeep Bhattacharya^{d,e,*}

^a Department of Computing, Goldsmiths, University of London, SE14 6NW, UK

^b Laboratory of Neurobiology, University College London, WC1E 6BT, UK

^c The Institute of Music Physiology and Musicians' Medicine, Hanover University of Music and Drama, 30161 Hanover, Germany

^d Department of Psychology, Goldsmiths, University of London, SE14 6NW, UK

^e Commission for Scientific Visualization, Austrian Academy of Sciences, Vienna, 1220, Austria

ARTICLE INFO

Article history:

Received 19 June 2009

Revised 21 November 2009

Accepted 3 December 2009

Available online 11 December 2009

ABSTRACT

The ability to anticipate forthcoming events has clear evolutionary advantages, and predictive successes or failures often entail significant psychological and physiological consequences. In music perception, the confirmation and violation of expectations are critical to the communication of emotion and aesthetic effects of a composition. Neuroscientific research on musical expectations has focused on harmony. Although harmony is important in Western tonal styles, other musical traditions, emphasizing pitch and melody, have been rather neglected. In this study, we investigated melodic pitch expectations elicited by ecologically valid musical stimuli by drawing together computational, behavioural, and electrophysiological evidence. Unlike rule-based models, our computational model acquires knowledge through unsupervised statistical learning of sequential structure in music and uses this knowledge to estimate the conditional probability (and information content) of musical notes. Unlike previous behavioural paradigms that interrupt a stimulus, we devised a new paradigm for studying auditory expectation without compromising ecological validity. A strong negative correlation was found between the probability of notes predicted by our model and the subjectively perceived degree of expectedness. Our electrophysiological results showed that low-probability notes, as compared to high-probability notes, elicited a larger (i) negative ERP component at a late time period (400–450 ms), (ii) beta band (14–30 Hz) oscillation over the parietal lobe, and (iii) long-range phase synchronization between multiple brain regions. Altogether, the study demonstrated that statistical learning produces information-theoretic descriptions of musical notes that are proportional to their perceived expectedness and are associated with characteristic patterns of neural activity.

© 2009 Elsevier Inc. All rights reserved.

Introduction

The brain's ability to anticipate forthcoming events accurately and efficiently has a clear adaptive value and predictive successes and failures often entail significant psychological and physiological consequences, modulating arousal and affecting reward circuits in the brain (Schultz et al., 1997; Steinbeis et al., 2006). Both confirmation and violation of expectation are thought to be critical to the way in which a piece of music communicates emotion and creates aesthetic experiences of, for example, tension, disappointment, pleasure, humour, and frisson (Huron, 2006; Meyer, 1956; Narmour, 1990).

Previous research has investigated event-related potential (ERP) responses to violations of expectation in harmony (Koelsch et al., 2000, 2002b, 2008; Leino et al., 2007; Loui et al., 2005; Patel et al.,

1998; Steinbeis et al., 2006). Two characteristic brain responses are reported: the early (right) anterior negativity E(R)AN with a latency of 150–280 ms, and a later bilateral or right-lateralised negativity (N5) with a latency of 500 ms (Koelsch et al., 2000; Loui et al., 2005). The E(R)AN is thought to reflect the violation of harmonic expectation, while the N5 reflects a higher processing effort needed to integrate unexpected harmonies into the ongoing context (Steinbeis et al., 2006).

While harmony is important in Western tonal music, it plays a less significant role in other musical traditions, which emphasize pitch, timbre, and rhythm. To date, little is known about the neural correlates of expectation in these musical dimensions. However, there is a sparse literature reporting ERP responses to violations of melodic expectation, and the picture appears to be somewhat more complex than for violations of harmonic expectation. Early studies (Besson and Faita, 1995; Nittono et al., 2000; Paller et al., 1992; Schon and Besson, 2005; Verleger, 1990) focused on a late positive component (LPC) peaking between 300 and 600 ms at central and posterior sites. The amplitude and latency of the LPC are sensitive to

* Corresponding author. Department of Psychology, Goldsmiths, University of London, New Cross, SE14 6NW London, UK. Fax: +44 20 7919 7873.

E-mail address: j.bhattacharya@gold.ac.uk (J. Bhattacharya).

musical expertise, the familiarity of the melody, the degree of unexpectedness (Besson and Faita, 1995), and also to the timing of the unexpected note (Nitto et al., 2000). In a more recent study, Miranda and Ullman (2007) describe a functional dissociation between two ERP components: first, an early (150–270 ms) anterior-central negativity associated with out-of-key violations in both familiar and unfamiliar melodies; second, a posterior negativity in the 220–380 ms latency range elicited by both in-key and out-of-key violations of familiar melodies only. They suggested that these two components are driven by violations of musical rules (of tonality/harmony) and of veridical memory representations of familiar melodies, respectively.

Although these studies do reveal several neural mechanisms of melodic expectations, a few crucial limitations exist across them. The ecological validity is low since the stimuli have been artificially constructed or altered to produce expected and unexpected continuations of the melodic excerpt. This means that the results do not necessarily generalise to natural perception since the participants responded to stimuli that they are less likely to encounter in the natural environment (Clarke, 2005; Dowling, 1989). Further, these studies typically examined responses to penultimate or final notes (Miranda and Ullman, 2007 being a recent exception), which elicit patterns of expectation specifically related to tonal closure (Aarden, 2003). Furthermore, only ERPs were analysed, so the studies do not reveal any information about oscillatory neuronal synchronization, within and between cortical regions, yet there is widespread evidence that neuronal synchronization, both local and global, acts as a flexible

mechanism of attentional and emotional selection (Bhattacharya et al., 2001; Fries, 2005; Varela et al., 2001). Both of these are crucial in musical expectations.

To address these issues, we systematically investigated melodic expectation using a tripartite approach involving distinct computational, behavioural, and electrophysiological (electroencephalogram, EEG) components. Instead of designing stimuli to match experimental hypotheses, our work started with an operational definition of musical expectation embodied in a computational model (Pearce and Wiggins, 2006) that views the brain as a learning system whose goal is to predict future events as accurately as possible (Friston, 2005; Schultz et al., 1997). We considered natural musical stimuli (melodies from English hymns encoded as MIDI files) and used the model to identify notes which are highly probable (with high probability) and highly improbable (with low probability) in context. This not only provided a more ecologically valid paradigm but also allowed us to test the predictions of the model explicitly, while the choice of high- and low-probability notes was intended to maximise any electrophysiological effects. In the behavioural experiment, participants performed a visually cued auditory expectation task where a response was made without pausing the melody, thereby avoiding the problem of spurious perception of closure (Figs. 1A and B). In the electrophysiological experiment, no explicit instruction was given, and the participants listened to the melodic excerpts with eyes closed. Since the participants were not explicitly informed about the locations of the probed notes and also their probabilities, the differences in neural responses between the two types of notes are likely to reflect the

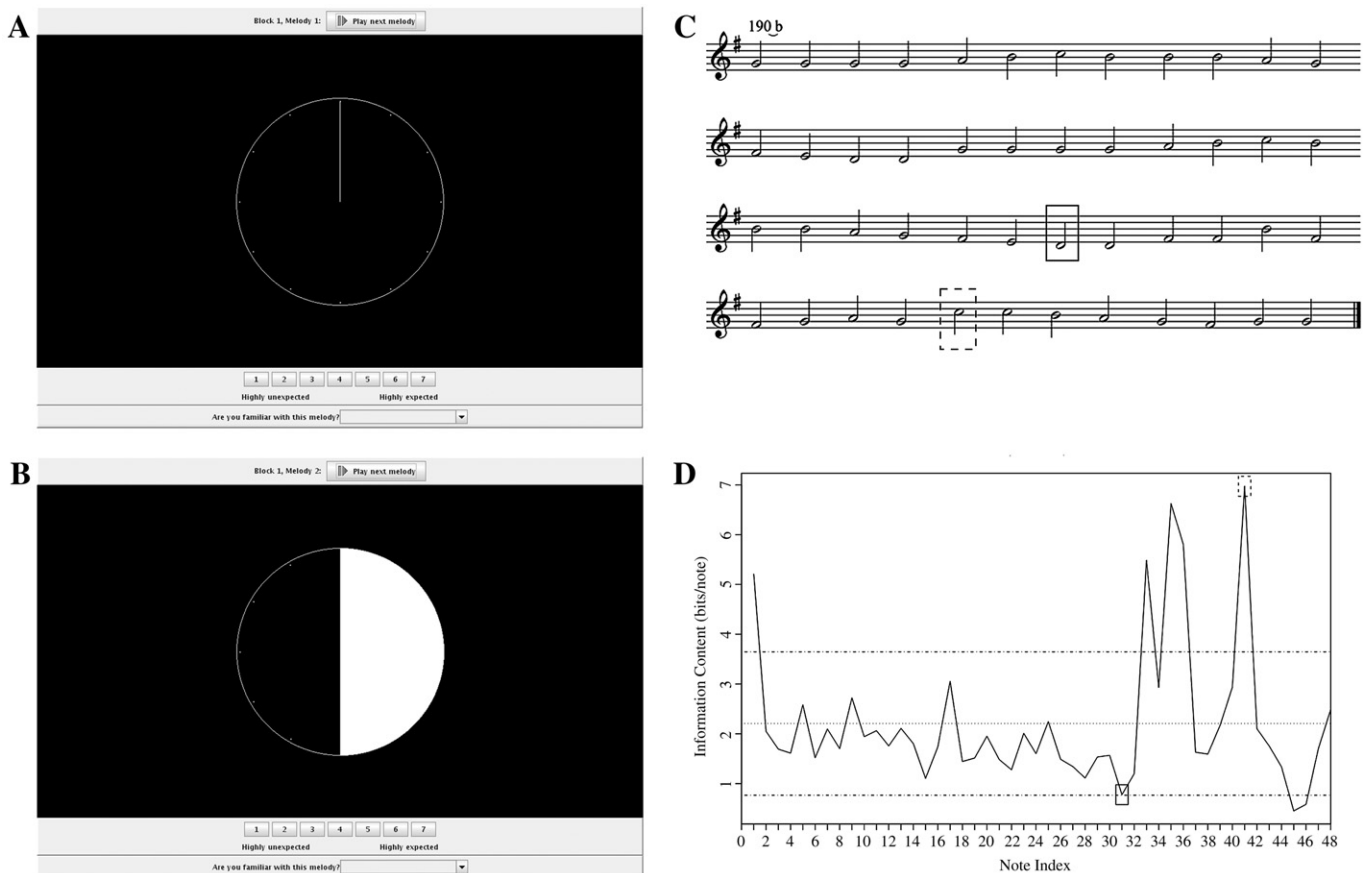


Fig. 1. The behavioural paradigm, stimuli and model predictions. (A) The display used to present the stimuli and collect behavioural responses. (B) As the melody approaches a probed note, the hand of the clock moves clockwise (shown here at 6:00); the participants were instructed to respond to the note whose onset coincides with the hand reaching 12:00. (C) One of the hymn melodies used as stimuli: Hymn 190. The boxes indicate the probed notes in the low-probability (dashed line) and high-probability (filled line) conditions. The high- and low-probability probe notes were associated with low and high information content respectively as computed by the model. (D) An information content profile representing the expectations of the model for each note in Hymn 190. The boxes show the low-probability (dashed line) and high-probability (filled line) notes. The horizontal lines indicate the mean and standard deviation of the information content in the melody.

implicit nature of melodic pitch expectations. EEG signals were analysed in terms of ERP, neuronal oscillatory activity within, and phase synchronization between near and distant cortical regions across the brain. For the behavioural experiment, we hypothesized that notes predicted by the model, high- and low-probability notes, would be subjectively perceived as proportionally more and less expected, respectively. For the electrophysiological experiment, we hypothesized that processing of high- and low-probability notes would be characterised by distinct patterns of brain responses.

Materials and methods

Computational models of musical expectation

Existing computational models of musical expectation fall into two groups: (i) supervised or rule-based models which generate expectations according to some static rules that predict what will happen next in a given context and (ii) unsupervised models which generate expectation based on learned associations between events that co-occur and uses these acquired associations to predict future events on the basis of the current context.

Probably the best-known rule-based account of melodic expectation is that of Narmour (1990) which proposes that expectations are partly based on a number of so-called *bottom-up* principles, inspired by the Gestalt principles in vision, that are universal properties of the human auditory perceptual system. These principles make predictions about the size of a melodic interval (the difference in pitch between two consecutive notes) given the preceding interval. While Narmour's theory is presented in a music-theoretic way, its principles have been quantified and implemented by psychologists attempting to test its predictions. The best-performing and most parsimonious implementation is known as the *two-factor* model (Schellenberg, 1997). It consists of two principles, *pitch proximity* and *pitch reversal*: Pitch proximity reflects an expectation for small melodic intervals (i.e., two consecutive notes are similar in pitch), pitch reversal embodies and expectation for change of pitch contour (e.g., a descending pitch interval after a rising one) following large melodic intervals.

In this study, we adopted a model of musical expectation based on statistical learning, probability estimation and information theory. We hypothesised that, while listening to music (or, indeed, perceiving other phenomena which are sequential in time), the brain anticipates or predicts possible continuations of the current (musical) context. These predictions were based on a model of the perceived domain (music, in the current case) formed by an inductive process of unsupervised statistical learning of perceived sequential structure. The learned model encodes past experience, and can be used to anticipate future events on that basis, using its acquired statistical knowledge of sequential structure to generate estimates of the probabilities of known events occurring, conditional upon the current sequential context. In music, such expectations depend on many aspects of musical structure, including harmony, but here we focused on pitch expectations for single note continuations to melodic contexts. Specifically, we predicted that a listener estimates the probability of different anticipated pitched continuations to a melody using the frequency with which each one has followed the context in his/her previous musical experience. High-probability notes are expected, while low-probability notes are unexpected.

We have developed a computational model that embodies this account of expectation. The model's goal is to estimate in any context a conditional probability distribution governing the probability of the pitch of the next note in a melody given the preceding notes. Thus, if we represent a melody X of n notes as a sequence of pitches, x_1, x_2, \dots, x_n , the goal of the model is to estimate the conditional probability of the i -th note in the melody, $p(x_i|x_1, \dots, x_{i-1})$. Given these estimates of conditional probability, the model's expectations may be quantified by *information content* (MacKay, 2003). The information content of

the i -th note in a melody is simply the negative logarithm, base 2, of the probability: $-\log_2(p(x_i|x_1, \dots, x_{i-1}))$. The information content of an event is inversely proportional to its probability (i.e., it reflects unexpectedness) but has a more convenient scale (probabilities often become vanishingly small) and its units have a natural interpretation in information theory as the minimum number of bits required to encode the event in context (Shannon, 1948).

The model has been designed to produce probability estimates that are as accurate as possible and we now summarise how this is achieved. Probabilities were estimated using n -gram models commonly used in statistical language modeling (Manning and Schütze, 1999). An n -gram is a sequence of n symbols and an n -gram model is simply a collection of such sequences each of which is associated with a frequency count. During the *training* of the statistical model, these counts were acquired through an analysis of some corpus of sequences (the training set) in the target domain. When the trained model is exposed to a sequence drawn from the target domain, it uses the frequency counts associated with n -grams to estimate a probability distribution governing the identity of the next symbol in the sequence given the $n-1$ preceding symbols. The quantity $n-1$ is known as the *order* of the model and represents the number of symbols making up the context within which a prediction is made.

The most elementary n -gram model of melodic pitch structure (a monogram model where $n=1$) simply tabulates the frequency of occurrence for each chromatic pitch encountered in a traversal of each melody in the training set. During prediction, the expectations of the model are governed by a zeroth-order pitch distribution derived from the frequency counts and do not depend on the preceding context of the melody. In a digram model (where $n=2$), however, frequency counts are maintained for sequences of two pitch symbols and predictions are governed by a first-order pitch distribution derived from the frequency counts associated with only those digrams whose initial pitch symbol matches the final pitch symbol in the melodic context. Fixed-order models such as these suffer from a number of problems. Low-order models (such as the monogram model discussed above) clearly fail to provide an adequate account of the structural influence of the context on expectations. However, increasing the order can prevent the model from capturing much of the statistical regularity present in the training set. An extreme case occurs when the model encounters an n -gram that does not appear in the training set in which case it returns an estimated probability of zero. To address these problems, the models used in the present research maintain frequency counts during training for n -grams of all possible values of n in any given context. During prediction, distributions are estimated using a weighted linear combination of all models below a variable order bound, which is determined in each predictive context using simple heuristics designed to minimize model uncertainty. The combination is designed such that higher-order predictions (which are more specific to the context) receive greater weighting than lower-order predictions (which are more general). In a given melodic context, therefore, the predictions of the model may reflect the influence of both the digram model and (to a lesser extent) the monogram model discussed above. Furthermore, in addition to the general, low-order statistical regularities captured by these models, the predictions of the model can also reflect higher-order regularities which are more specific to the current melodic context (to the extent that these exist in the training set). Pearce and Wiggins (2004) give a comprehensive account of the generation of predictions from the trained models, the details of which lie beyond the scope of the present article. One final issue to be covered regards the manner in which the statistical model exploits the representation of multiple features of the musical surface described above. The modeling process begins with the selection of a set of features of interest and the training of distinct n -gram models for each of these features.

Finally, for each event in a melody, each feature is predicted using two models: first, a *long-term* model that was trained over the entire training set and encodes learnt schematic expectations based on long-term exposure to a large corpus of music; and second, a *short-term* model that is trained incrementally for each individual melody and encodes local expectations learnt incrementally by analysis of that melody during (simulated) listening. The probability distributions returned by these two models are combined by taking the product of the weighted probability estimates returned by each model for each possible value of the pitch of the next event and then normalizing such that the combined estimates sum to unity over the pitch alphabet. Greater weights are assigned to models whose predictions are associated with lower entropy (or uncertainty) at that point in the melody. The use of long- and short-term models is intended to reflect the influences on perceptual expectations of both long-term musical experience and incrementally increasing knowledge of the structure of the piece of music one is currently listening to. The entropy-based weighting method and the use of a multiplicative (as opposed to a linear) combination scheme both improve the performance of the model in predicting unseen melodies (Pearce et al., 2005; Pearce and Wiggins, 2004, to which the reader is referred for a full technical description of the combination of predictions).

For the purposes of this study, the model derives its pitch predictions from a representation of pitch interval and scale degree reflecting the fundamental influence of melodic and tonal structure respectively (though in other work we use richer representations). Each note in a melody is represented by a pair of values: first, the pitch interval preceding the note; and second, the scale degree of the note relative to the notated key of the melody. The long- and short-term models produce probability distributions generated over an alphabet of such pairs and these are converted into probabilities for concrete chromatic pitches before being combined. The long-term component was trained on the corpus of melodies shown in Table 1, which are intended (subject to practical constraints) to represent the long-term musical experience of a listener familiar with the melodic style of western tonal music (of which the English hymns used as stimuli are another example). Fig. 1C shows an example of a melodic excerpt, and its information content profile, as estimated by our model, is shown in Fig. 1D.

Ethics statement

Both behavioural and electrophysiological experiments were approved by the local ethics committee of the Department of Psychology at Goldsmiths College, University of London. Informed written consent was obtained from all participants.

Behavioural experiment

Participants and experimental design

Forty participants (17 females and 23 males, age range 19–72 years, mean age 27.58 years) consisting of 20 musicians (10 females and 10 males, age range 19–72 years, mean age 30.5 years, 18 right-handed, 2 left-handed) and 20 non-musicians (7 females and 13 males, age range 19–40 years, mean age 29.6 years, 17 right-handed, 3 left-handed) took part in the experiment. Musicians had an average of 12.5 years of training and had played a musical instrument for an average of 22.5 years, whereas non-musicians had an average of

0.48 years of formal training and had played an instrument for an average of 1.6 years. All participants were either students or staff at Goldsmiths, University of London, and were in good health, with normal hearing and no past history of neurological illness. In total, five participants self-identified as being left-handed.

The stimuli consisted of 28 hymn melodies (see Table S1 for a list) selected and transcribed from a Church of England hymnal (Nicholson et al., 1950). The melodies were played untransposed in their original keys (a range of major keys except for one hymn which was in D minor). All of them were notated with a time signature of 4/2 and began on the first beat. All the phrases within each melody were of equal duration. In order to focus specifically on pitch expectations, rhythmic structure was removed in a musically sensitive manner by a skilled musicologist, so that each note had the same duration and equivalent interonset interval of 1 s (120 bpm). Specifically, dotted rhythms were evened out, pairs of crotchets were replaced by minims on the harmony note and long notes (usually phrase-final semi-breves) were replaced by pairs of minims. This was done to avoid the potentially confounding effects of expectations regarding the timing of notes, which is known to influence neural responses to manipulations of melodic expectation (Nitto et al., 2000). In contrast to other studies (e.g., Miranda and Ullman, 2007), no notes were added nor were existing notes modified to maintain the identity of the melody and to preserve the ecological validity. The melodies were rendered as MIDI files using the *Electronic Piano 1* instrument (program 5) of a Roland Sound Canvas (SC-88) MIDI synthesiser. Although there was no expressive timing, the dynamics were subtly varied so as to introduce a pattern of accents that emphasised the notated metre.

In each melodic excerpt, two notes were selected as locations to probe the expectations of listeners. The probe locations were selected according to the predictions of the computational model of perceived pitch expectations in melody (Pearce and Wiggins, 2006) described earlier. According to the model, in the melodic context in which they appear, one of these notes has a high conditional probability of occurrence while the other has a low probability of occurrence. Fig. 2A shows the distribution of information contents of notes in the entire collection of hymns, while Fig. 2B shows the same distribution for the selected probe notes. The bimodal distribution evident for the probe notes reflects the selection of notes with high and low conditional probability. The probe locations were positioned at least 8 notes after the beginning of the melody and 8 notes after the previous probe note to give the participants some context for their expectation judgements and allow them to reorient attention to the melody after one response before having to respond again. Fig. 1C shows a melody with the probed notes indicated. Each musical excerpt lasted for less than one minute (35–40 s).

Participants were instructed to listen carefully to the musical stimuli presented binaurally by headphones. For each stimulus, the probe locations were indicated by the rotating hand of a clock, which counted down, stepwise, in quarters, in time with the music, informing the participant in advance when they were required to respond. The participant was required to give a rating on a Likert scale of 1 to 7 (1 being highly unexpected and 7 being highly expected) on how expected or unexpected the probe note was in the context of the preceding melodic passage. After listening to each melody, the participants were asked to indicate if it was familiar to them. Practice trials were provided for familiarisation with the experimental procedure. The order of presentation of the stimuli was randomised across participants.

Electrophysiological experiment

Participants and experimental design

Twenty healthy adult humans (13 males and 7 females, age range 19–26 years, mean age 20.7 years) participated in the EEG study. None of the participants had taken part in the behavioural study. All

Table 1
The melodic datasets used for training the computational model.

Description	No. compositions	No. notes	Mean length
Canadian folk songs	152	8553	56.27
Chorale melodies	185	9227	49.88
German folk songs	566	33,087	58.46
Total	903	50,867	56.33

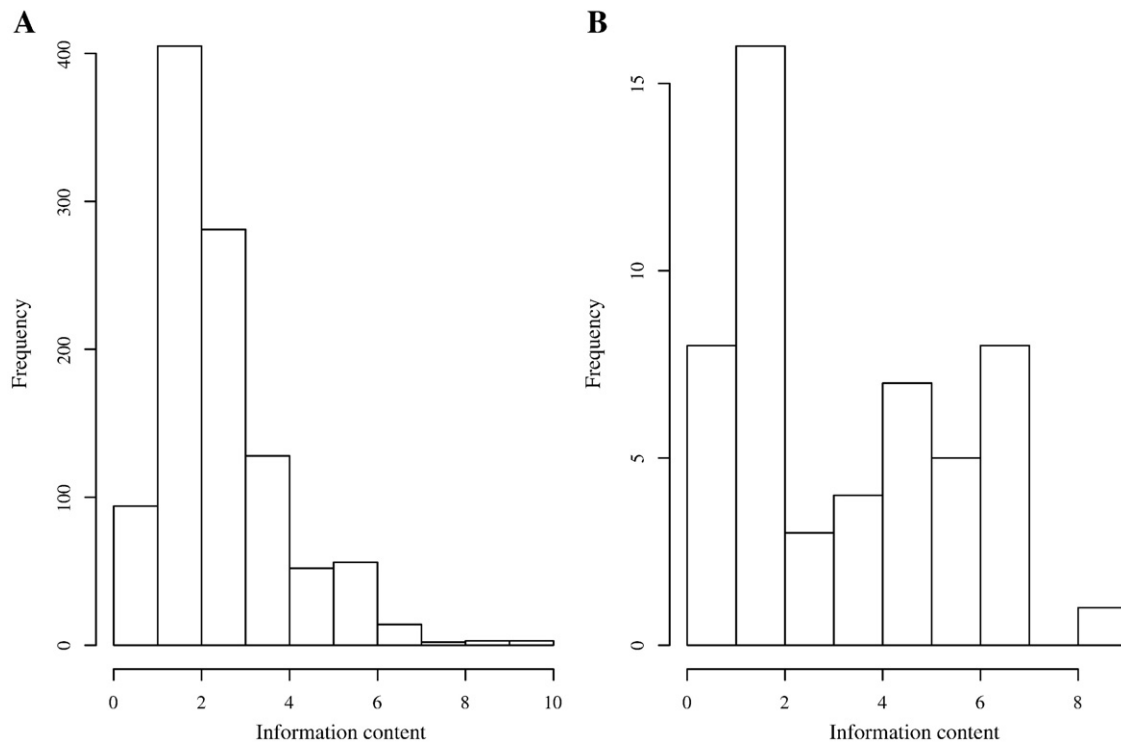


Fig. 2. Distribution of note information contents. (A) The distribution of all notes across melodies. (B) The distribution for the selected probe notes (both high-probability and low-probability conditions).

participants were in good health, had no past history of neurological disorders, and had no reported hearing difficulties. None of the participants reported having any formal musical training.

The same set of 28 melodic excerpts selected for the behavioural experiment was used here. To avoid artefacts caused by eye/head movements, the participants were asked to listen attentively to each melodic excerpt with eyes closed. No explicit expectedness ratings were requested, and the participants were not made overtly aware of the location of the probe notes, thereby emphasizing the implicit aspect of melodic processing.

Data acquisition and preprocessing

EEG signals were recorded from 28 Ag/AgCl electrodes according to the extended 10–20 system (Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, C5, C3, Cz, C4, C6, CP5, CP3, CPz, CP4, CP6, P7, P3, Pz, P4, P8, O1, Oz, O2) (Jasper, 1958). The vertical and horizontal electro-oculograms (EOG) were bipolarly recorded from four additional channels to monitor eye movements and blinks. Sampling rate was 500 Hz. The recorded EEG signals were algebraically re-referenced against the mean amplitude of two mastoids. All electrode impedances were kept below 5 kΩ.

We used the EEGLAB Matlab® Toolbox (Delorme and Makeig, 2004) for visualization and filtering purposes. A high-pass filter at 0.5 Hz was applied to remove linear trends and a notch filter at 50 Hz (49–51 Hz) was applied to eliminate line noise. The EEG data were further cleaned of remaining artefacts by means of wavelet-enhanced independent component analysis (Castellanos and Makarov, 2006), after first computing the ICA components with the FastICA algorithm (Hyvarinen and Oja, 2000). The data epochs representing single experimental trial time-locked to the onset of the probe notes were extracted from –1000 ms to 1000 ms.

Data analyses

We performed the following types of data analysis. (i) The standard time-averaging technique to analyze the ERPs associated with high- and low-probability notes. The ERPs for each subject and

condition were baseline-corrected with the mean activity from 200 to 0 ms before the note onset. Next, we computed the wavelet based time–frequency representations (TFR) to analyze (ii) the spectral power of the oscillatory contents and (iii) the spatiotemporal dynamics of the phase coupling as measured by bivariate synchronization analysis (Lachaux et al., 1999; Pereda et al., 2005) and by the phase lag index (Stam et al., 2007) for investigating phase synchrony.

A complex Morlet wavelet was used to extract time–frequency complex phases, at an electrode i and epoch k , and amplitudes of the EEG signal $x(t)$. The frequency domain was sampled from 2 to 60 Hz with a 1-Hz interval between each frequency. To study changes in the spectral power, we used the TFR of the wavelet energy (Tallon-Baudry et al., 1997). After removing the baseline level (200 prestimulus), we normalized the wavelet energy with the standard deviation of the baseline period and expressed it as percentage of power change. Oscillatory activity was analyzed in the theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (31–60 Hz) frequency bands.

Bivariate phase synchronization is a useful approach to assess phase synchronization between neurophysiological signals (Hurtado et al., 2004; Lachaux et al., 1999; Pereda et al., 2005; Rodriguez et al., 1999; Tass et al., 1998). At a centre frequency f , the strength of the phase coupling between two electrodes i and j at time t was computed as:

$$\bar{R}_{ij} = \left| \frac{1}{n} \sum_{k=1}^n \exp(i(\phi_{jk} - \phi_{ik})) \right|,$$

where n is the number of epochs. This index approaches 0 (1) for no (strict) phase relationship between the considered electrode pair across the epochs. The average of this index across pairs of electrodes represents a measure of global synchronization strength (\bar{R}). For the bivariate synchronization analysis, a modified version of the nearest-neighbour Hjorth Laplacian algorithm computed by Taylor's series expansion (Lagerlund et al., 1995) was applied to avoid the spurious increase in correlations introduced by the common reference and the volume conduction (Nunez et al., 1997).

The pairwise synchronization index \bar{R}_{ij} , albeit a powerful indicator of phase coupling, cannot dissociate between zero phase lag from constant phase lag. Therefore, we also calculated the phase lag index (PLI), a novel measure of detecting phase lag (delayed) synchronization. This measure is shown to be less affected by the influences of common sources and active reference electrodes (Stam et al., 2007). The PLI was computed from the time series of the phase differences of two signals at epoch k , according to:

$$PLI_{ij}(t) = \left| \frac{1}{n} \sum_{k=1}^n (\varphi_{jk}(t) - \varphi_{ik}(t)) \right|$$

The PLI ranges between 0 for no coupling or coupling around 0 mod π , and 1 for nonzero phase coupling. We used an average reference before computing the PLI.

At each frequency from 2 to 60 Hz with a step size of 1 Hz, the indexes \bar{R}_{ij} and PLI_{ij} were computed and baseline-corrected (baseline being 200 ms prestimulus). They were subsequently averaged across electrodes to obtain a measure of the global synchronization strength, \bar{R} and \overline{PLI} . We focused our analysis on the theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (31–60 Hz) frequency bands.

Statistics

To assess the statistical differences in the spectral power and phase synchronization and phase lag indices, we first averaged these measures for each participant and condition across all electrodes. Next, for each time–frequency point in the bands under study, the averaged measures were analyzed by means of a nonparametric pairwise permutation test (Good, 2005). Similarly, the ERP values for each condition and participant were averaged across all electrodes over the regions of interest, which were selected by visual inspection of the topographical distribution of the ERP waveforms (see Results). The evaluation of the statistical differences in the grand-averaged ERPs was also performed by means of a pairwise permutation test. Permutation tests are usually recommended when the distributions are not necessarily Gaussian. Also, the permutation test provides exact significance levels even for small sample sizes and small differences between conditions. For all indices, the test statistic was the difference between the two sample means: low-probability notes minus high-probability notes. For all statistical comparisons, the permutation tests were computed by first generating 5000 random permutations of the joint sample, then taking the first and second half of each randomly generated joint distribution and calculating replications of the test statistics. The p values were then obtained as the frequencies that the replications had larger absolute values than the experimental difference. When multiple permutation tests were performed, such as the case of several frequencies or ROIs, Bonferroni correction was applied to keep the family-wise error rate at 0.05.

Results

Behavioural experiment

There were two categories of probe notes: high- and low-probability. The size of the pitch interval preceding the high-probability notes (mean = 2.4 semitones) was found to be smaller than that

preceding the low-probability notes (mean = 5.3, $t = 6.6$, $p < 0.01$). Furthermore, using the empirical key profiles of Krumhansl and Kessler (1982) derived from the judgements of expert musicians, the high-probability notes were found to be more stable within the notated key of the melody than the low-probability notes, $t = -6.04$, $p < 0.01$.

The mean expectedness ratings and response times are summarised in Table 2 and shown graphically in Fig. 3. The ratings and response times show a small but significant negative correlation, $r = -0.24$, $p < 0.01$, indicating that participants responded faster to more expected notes. Fig. 3 suggests that high-probability notes are perceived as more expected and elicit faster responses than low-probability notes and also that both training and familiarity increase expectedness and reduce response times. To investigate the statistical significance of these effects, separate $2 \times 2 \times 2$ mixed-factorial ANOVAs with three factors—group (two levels: musicians vs non-musicians), familiarity (two levels: familiar and unfamiliar), and probe type (two levels: low probability vs high probability)—were applied to both ratings and response times.

For the perceived expectedness ratings, the analysis revealed significant main effects of probe type, $F(1,2050) = 245.40$, $p < 0.01$, of musical training, $F(1,2050) = 62.03$, $p < 0.01$, and of familiarity, $F(1,2050) = 53.51$, $p < 0.01$. There were also significant two-way interactions between training and familiarity, $F(1,2050) = 10.96$, $p < 0.01$, training and probe type, $F(1,2050) = 6.31$, $p < 0.05$, as well as probe type and familiarity, $F(1,2050) = 24.16$, $p < 0.01$. The three-way interaction between probe type, training, and familiarity was also significant, $F(1,2050) = 6.47$, $p < 0.05$. The analysis of the response times revealed significant main effects of probe type, $F(1,2050) = 17.27$, $p < 0.01$, musical training, $F(1,2050) = 21.14$, $p < 0.01$, but not familiarity. There were also significant interactions between training and familiarity, $F(1,2050) = 12.35$, $p < 0.01$, and between training and probe, $F(1,2050) = 12.6$, $p < 0.01$. No other interactions were significant.

To examine the effects of musical training in more detail, the number of years of training and instrumental experience were regressed, together with the dummy predictors probe type and familiarity, on the responses. The resulting model accounted for approximately 37% of the variance in the responses ($R^2 = 0.37$, $R^2_{\text{adj}} = 0.37$, $F(4,2055) = 32.9$, $p < 0.01$) with all predictors making a significant contribution to the model ($p < 0.01$) except years of instrumental experience. This indicates that expectations are affected linearly by years of training but not years of instrumental experience. A similar multiple regression analysis for the reaction times (including probe type but not familiarity as a predictor) yielded a model that accounted for 7% of the variance ($R^2 = 0.07$, $R^2_{\text{adj}} = 0.07$, $F(3,2056) = 517$, $p < 0.01$), but neither years of training nor years of instrumental experience made a significant contribution to the model.

To investigate further the correspondence between the information content of the probe notes as estimated by the computational model and the expectedness ratings made by the participants, we carried out two correlation analyses. The expectedness ratings and responses times were first averaged across participants for each note within each melodic excerpt. The averaging is warranted by high interindividual consistency for the expectedness ratings (Cronbach's $\alpha = 0.91$) and for the response times ($\alpha = 0.95$). The analysis yielded high correlations in the predicted directions between the information

Table 2
Mean expectedness ratings and response times.

			Expectedness ratings		Response times (ms)	
			Musicians	Non-musicians	Musicians	Non-musicians
Familiar melodies	Type of note	Low probability	5.98	4.46	2144	2710
		High probability	6.83	5.94	1357	2314
Unfamiliar melodies	Type of note	Low probability	4.09	3.77	2596	2505
		High probability	6.50	5.75	1729	2065

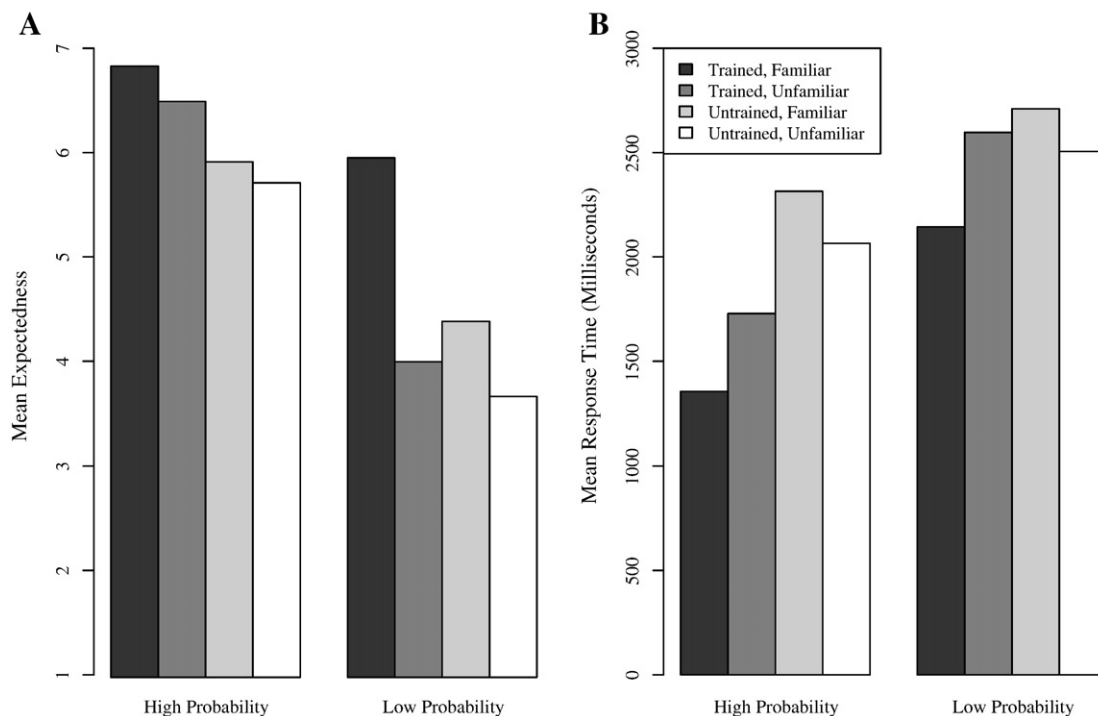


Fig. 3. The behavioural data. (A) The mean expectedness responses and (B) the mean response times for the different conditions of high and low probability, musical training and familiarity.

content of the model and both the mean expectedness ratings, $r = -0.89$, $p < 0.01$, and mean response times, $r = 0.75$, $p < 0.01$, of the participants. Therefore, the model accounts for ~78% of the variance in the ratings and ~56% of the variance in the response times.

Finally, to assess the significance of these results, we compared the performance of our proposed computational model with a competing rule-based model, the two-factor model of Schellenberg (1997). A multiple regression analysis was conducted in which the two predictors of the two-factor model, proximity and reversal, were regressed on both the mean expectedness ratings and the response times. The two-factor model accounted for approximately 56% of the variance in the ratings ($R^2 = 0.57$, $R^2_{\text{adj}} = 0.56$, $F(2,49) = 32.9$, $p < 0.01$) and 33% of the variance in the response times ($R^2 = 0.36$, $R^2_{\text{adj}} = 0.33$, $F(2,49) = 13.7$, $p < 0.01$). At an α level of 0.5, only proximity made a significant contribution to the fit of the two-factor model. The fit of the two-factor model to the data can be compared to the fit of the statistical model using Williams' t statistic for comparing dependent correlations between non-nested models regressed on the same dependent variable (Steiger, 1980). This analysis indicated a significant difference for the ratings, $t(49) = 2.5$, $p < 0.05$, and a marginally nonsignificant difference, $t(49) = 1.9$, $p = 0.06$ for the response times.

ERP analysis

First we compared event-related-potentials elicited by low-probability notes with high-probability notes. Fig. 4 shows ERPs at selected frontal, central, and fronto-central electrode positions. Both notes elicited a sharp negative peak at 100 ms latency (N1 component). A robust difference was found in the latency range 300–600 ms, where ERPs of low-probability notes were associated with a reduced negativity as compared with ERPs of high-probability notes. However, this effect was reversed at centroparietal electrode, where low-probability notes were associated with more negative deflections. The topography of difference (low-probability – high-probability) ERP responses is shown in Fig. 5 at three different time windows. For the early time window (90–160 ms), an enhanced

negativity for low-probability notes was found at left centroparietal and midline electrodes. In contrast, for the later time windows (300–500 ms), low-probability notes elicited enhanced positivity over the right frontal region. Interestingly, another negative deflection was observed over the centroparietal regions between 400 and 500 ms.

Therefore, we selected two spatial regions of interest (ROIs) for the subsequent statistical analyses: the right frontal electrode region (FP2, F4, F8) and the left and mesial centroparietal region (C3, CP3, Cz, CPz). By means of the non-parametric permutation test, we looked for differences between the ERP waveforms of high and low-probability notes, between 0 and 600 ms and in the two mentioned ROIs. The statistical test yielded a significant difference ($p < 0.025$, Bonferroni-corrected; two ROIs) at the right frontal region from 280 to 600 ms and at the left-mesial centroparietal region from 400 to 450 ms. The first effect reflected the positive deflection from 280 to 600 ms in the difference ERP waveforms over right frontal regions, which was due to smaller negative ERP amplitude for low-probability than for high-probability notes. The second effect reflected the opposite phenomenon: the ERPs of low-probability notes over centroparietal regions from 450 to 550 ms had larger negative amplitudes than those of high-probability notes. Interestingly, the positive deflection from 280 to 600 ms over the right frontal region was observed in 19 out of 21 participants. The late negativity from 400 to 450 ms at left-mesial centroparietal electrodes appeared in 17 out of 21 participants. Consequently, the novel ERP findings reported here were highly consistent across subjects.

Time-frequency analysis

Next we calculated the time-frequency representations (TFR) of the oscillatory brain responses. Figs. 6A and B show the difference (low-probability – high-probability) time-frequency maps of the alpha (8–13 Hz) and the beta band (14–30 Hz) spectral power, averaged across all electrode locations, respectively. A permutation test was applied to find the time-frequency window of statistical significance between the two conditions. In the lower alpha band, a significant increase in the spectral power was observed during 700–

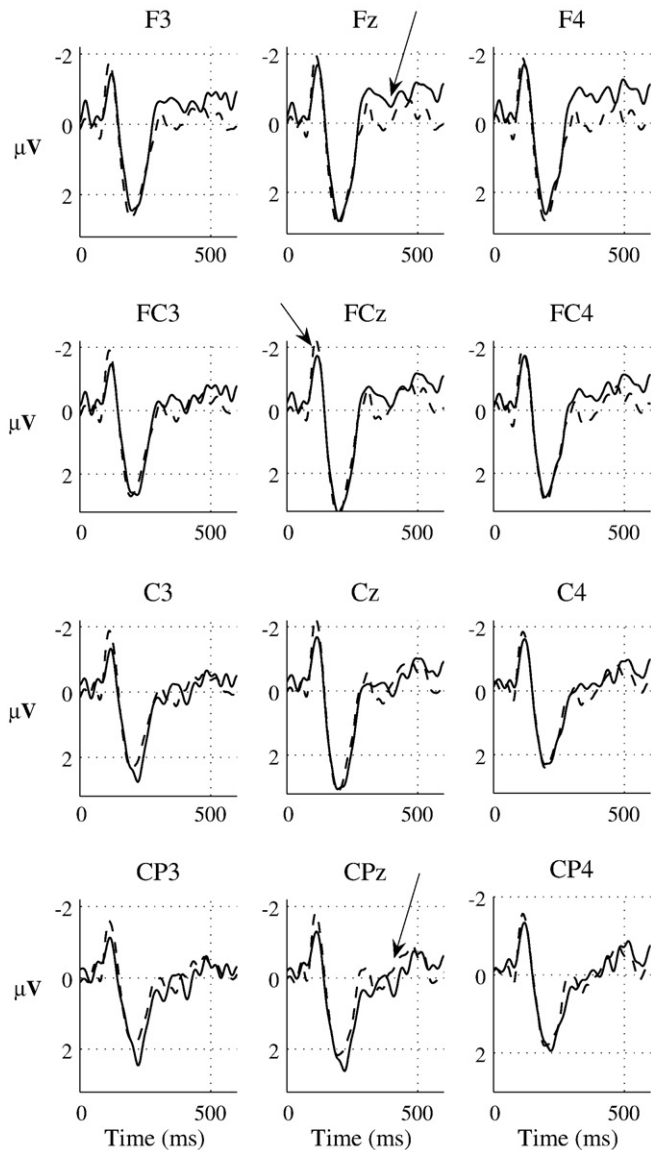


Fig. 4. ERPs to low- and high-probability notes. Grand-average ERPs elicited at frontocentral electrode sites by low-probability notes (dashed line) and high-probability notes (bold line). The onset of the probe notes was at 0 ms. The arrows indicate the late positive deflection in the difference ERP at electrode Fz, the early negativity at electrode FCz and the late negative deflection in the difference waveforms at electrode CPz.

900 ms ($p < 0.008$, Bonferroni-corrected); it was primarily due to a stronger decrease in the oscillatory activity at 7–8 Hz of high-probability notes as compared to low-probability notes. This effect

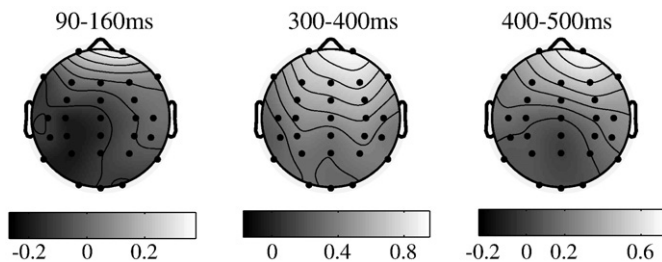


Fig. 5. Topographical maps of ERP differences. Isovoltage maps for the difference (low-probability – high-probability) ERP waveforms at three time windows: 90–60 ms (left), 300–400 ms (middle), and 400–450 ms (right). Within each time window, ERP amplitudes were averaged.

was spatially localized over left temporal regions (figure not shown). In the beta band, a significant increase in the spectral power was observed during 500–550 ms; it was due to both a strong burst of 18–23 Hz oscillations for low-probability notes and a minor decrease in 17–23 Hz oscillation for high-probability notes (see [Supplementary Figs. S1 and S2](#)). The spectral burst of beta oscillations for low-probability notes was localized over left mesial centroparietal areas (mainly C3, CP3, CPz, Pz; [Fig. 7C](#)).

Phase synchronization analysis

Both ERPs and TFRs represent locally synchronized neuronal responses, whereas neuronal assemblies communicating over long distances are revealed by methods based on phase synchrony. Therefore, we calculated the bivariate phase synchrony measure and phase lag index (see [Materials and methods](#)). The former measure is sensitive to synchrony with both zero and non-zero phase lag, whereas the latter is sensitive to only non-zero phase lag. The differences in phase synchrony measure between the two conditions were calculated across different frequency bands, and the results were most robust in the beta frequency band ([Fig. 7A](#)). We observed that the processing of low-probability notes was associated with a robust increase (with respect to their baselines) in the degree of global phase synchronization in the beta band and with a time span of 500–550 ms, whereas the processing of high-probability notes was associated with a decrease in the same time-frequency window. The permutation test yielded a statistically significant difference between conditions at 21–30 Hz and 500–550 ms ($p < 0.003$; Bonferroni-corrected). [Fig. 7B](#) shows the differences in global phase lag index (PLI) between low-probability and high-probability notes, showing a similar statistically significant effect at 20–22 Hz and 500–550 ms. There was an additional significant effect in the beta band in the global, PLI namely an increase during 100–150 ms. However, the later effect (500–550 ms) was more robust and common across both measures, so this time window was selected for further analysis of the topographical patterns of the pairwise phase coupling as revealed by both measures ([Figs. 7C and D](#)). In both cases, we selected a threshold of 50% of the maximum phase coupling for each measure, which produced adequate suprathreshold pairs. The spatial pattern based on bivariate phase synchrony reflected increases in long-range interhemispheric connections, and most coupled pairs were located over centroparietal and occipital regions. On the other hand, the spatial pattern based on phase lag index indicated a more heterogeneous pattern of pairs with PLI_{ij} increases/decreases but both predominantly over long-range connections.

Discussion

A tripartite approach was presented to study melodic pitch expectation by using computational, behavioural and electrophysiological methods. The study was anchored to a computational model, which acquired knowledge through unsupervised statistical learning of sequential structure in music and generates an estimation of the conditional probability (and information content) of musical notes. A high (inverse) correlation was found between the information content of a musical note and its perceived expectedness: as hypothesized, highly probable notes were perceived as expected while improbable notes were perceived as unexpected. Unlike previous studies, our novel behavioural paradigm allowed the participants to judge the expectedness of a note without interrupting the ongoing musical stimuli or artificially introducing tonal closure. In a separate electrophysiological experiment, we found that high- and low-probability notes were characterized by distinct patterns of neural activity, most predominantly in the beta band oscillations and phase synchronization. Altogether, these patterns suggest that

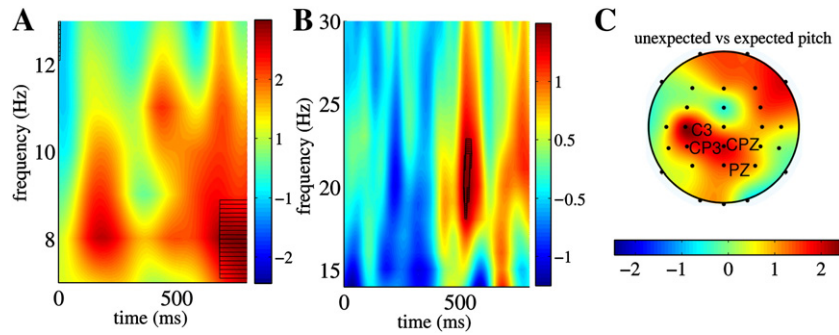


Fig. 6. Oscillatory responses to low- and high-probability notes. (A) Differences between the processing of low-probability and high-probability notes in the time course of the oscillatory activity of alpha band (8–13 Hz) averaged across all electrodes. (B) Same but in the beta band (14–30 Hz). Spectral power in the poststimulus period was expressed as standard deviations of the spectral power in the baseline (0–200 ms prestimulus). Regions of significant differences, as found by the permutation test, between conditions are denoted by the black contour. The significance levels for both alpha and beta bands were Bonferroni-corrected, yielding thresholds of 0.008 and 0.003, respectively. (C) Topographical map of the difference power at 18–23 Hz over the time window 500–550 ms. Bursts of beta oscillations were observed over the left and mesial centroparietal regions (C3, CP3, CPZ, Pz) during low-probability notes but were conspicuously absent during high-probability notes (see [Supplementary Figs. S1 and S2](#)).

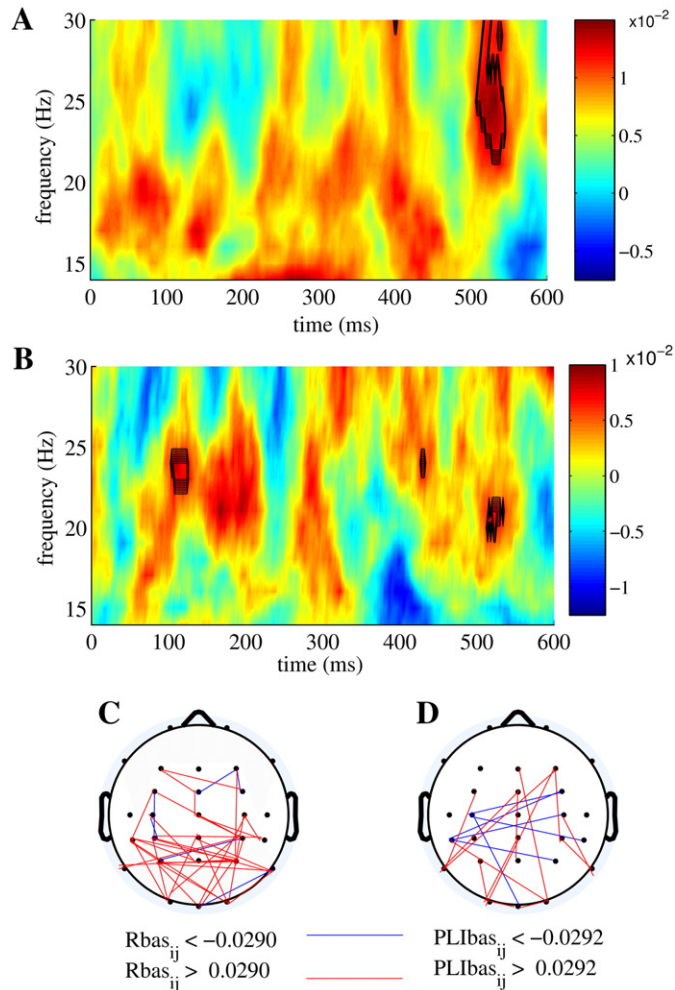


Fig. 7. Time course of phase coupling. (A, B) Time–frequency plots of the low-probability minus high-probability notes of the bivariate synchronization index R_{ij} (A) and the phase lag index PLI (B). Results were shown for beta band (14–30 Hz) and averaged across all pairs of electrodes. Black contours indicate regions of statistical significance (Bonferroni-corrected $p < 0.003$). (C, D) Topographical patterns showing the robust increases and decreases in phase coupling between two conditions using the bivariate synchronization measure (C) and the phase lag index (D). The threshold ($= 50\%$ of the maximal value for each measure) was chosen such that enough pairwise connections were visible. Note the increase in long-range and interhemispheric connectivity for low-probability notes.

melodic processing may involve an intimate coupling between the perception and production of sequential auditory events.

Behavioural support for the computational model

Our behavioural results strongly support our computational model: notes which it estimated to have high conditional probability (low information content) were rated as expected by listeners (and evoked faster responses) while those which it estimated to have low conditional probability (high information content) were rated as unexpected (and produced slower responses). Further, as in previous research ([Aarden, 2003](#); [Bharucha and Stoeckig, 1986](#)), musical notes that are unexpected in context lead to slower reaction times due to a longer integration process. The high-probability notes were found to be more tonally stable and preceded by smaller intervals than the low-probability notes. This corroborates previous reports that in actual music, tonally stable notes are more frequent than unstable ones ([Krumhansl, 1990](#)) and small intervals more frequent than large ones ([Huron, 2006](#)). According to the present account, statistical patterns such as these are learnt through exposure to music and influence perceptual expectations when listening to music. In support of this prediction, the expectations and response times of the listeners showed significant linear relationships with the degree of probability estimated by the model. Finally, the fit of the probabilistic model to the behavioural ratings was found to be better than that of a competing rule-based model ([Schellenberg, 1997](#)). However, the fact that large intervals are less probable (and hence, in the present account, less expected) than small ones remains a potential acoustic confound that should be addressed in future research, perhaps using artificially constructed stimuli or selecting low-probability notes that are preceded by a small interval in existing music.

Unsupervised statistical learning in music

The proposed approach emphasised the ability of the brain to learn statistical regularities in its perceptual input and used these regularities to predict future events. The theory was instantiated in a computational model ([Conklin and Witten, 1995](#); [Pearce et al., 2005](#); [Pearce and Wiggins, 2004](#)) which was based on n -gram methods often used in text compression ([Bell et al., 1990](#); [Bunton, 1997](#)) and statistical language modelling ([Manning and Schütze, 1999](#)). In contrast to rule-based models of cognitive processing (e.g., [Lerdahl and Jackendoff, 1983](#); [Narmour, 1990](#); [Temperley, 2001](#)) and nativist accounts of neural specialization for music ([Peretz, 2001; 2002](#); [Peretz and Hebert, 2000](#)), our approach predicted that neural responses should be sensitive to statistical predictability of the stimulus. Earlier, [Pearce and Wiggins \(2006\)](#) demonstrated that the information-

theoretic model predicts the melodic expectations of listeners better than existing rule-based models (Narmour, 1990; Schellenberg, 1997). Furthermore, in an fMRI study, Overath et al. (2007) demonstrated that activity in the planum temporale, a region of auditory association cortex, increases as a function of the information content (or predictability) of a melodic stimulus. Another point of divergence between our model and rule-based models is that it explicitly predicts and accounts for observed developmental changes in expectations through development (Schellenberg et al., 2002). Future research should focus on the neural basis of these developmental trajectories in the acquisition of auditory expectations. Finally, because it uses unsupervised learning and information-theoretic methods, the model generalises naturally to perceptual expectations in sequential cognitive domains beyond music. Indeed, cognitive difficulty of processing words corresponds to information-theoretic quantities such as information content (Levy, 2008) and entropy (Hale, 2006) while visual attention is drawn to probabilistically unexpected events in visual scenes (Itti and Baldi, 2006). Further research is needed to identify similarities and differences in the underlying neural processing involved in these different domains of perceptual processing.

Effects of musical training and familiarity

Since we included both musicians (musically trained participants) and non-musicians (without musical training) for our behavioural experiment, we could further investigate the effect of musical training on melodic expectations. We found that musicians gave higher ratings of expectedness and also responded faster than the non-musicians. These effects of musical training are perhaps not surprising, but they are interesting in the light of previous data suggesting that melodic expectation is not affected by formal training in music (Schellenberg, 1996). The significant interaction in both analyses indicates that, compared with the untrained participants, the musically trained participants showed a greater effect of increased expectedness, and corresponding decrease in reaction times, to the high probability, compared with the low probability, probes. That individuals with greater musical experience should find our stimuli more predictable is consistent with the learning-based account presented here.

We also examined the effects of familiarity with the melodies, which were mainly apparent in the expectedness ratings rather than the response times. The results demonstrated that expectations were higher for familiar melodies, that this was particularly so for the unexpected notes (an interaction between familiarity and probe type), and that both these effects were stronger for the musicians than for the non-musicians (a two-way interaction between familiarity and musical training and a three-way interaction between familiarity, musical training, and probe type). These findings corroborate previous results showing an influence of familiarity on both behavioural and neurophysiological indices of melodic expectation (Besson and Faita, 1995; Miranda and Ullman, 2007). Again, these effects of familiarity are consistent with the proposal that musical expectations are based on learning.

Novel ERP effects in response to melodic processing

To further validate the proposed information-theoretic model, we investigated the neural basis for the perceived difference between high-probability and low-probability musical notes in a melodic excerpt. We found distinct patterns of ERP responses, varying across both time and topography, associated with the low and high information content notes. Comparing the processing of these two types of notes, the main electrophysiological findings were (i) a positivity over the right frontal regions during 280–600 ms, which was mostly due to a larger negative amplitude in the ERPs of high-probability notes, and (ii) a negativity over the left-mesial centro-

parietal regions during 400–450 ms, which was due to a more negative ERP amplitude in response to low-probability notes. Furthermore, both significant effects were highly consistent across subjects (right frontal positivity observed in 19 participants and left-mesial centroparietal negativity observed in 17 participants). Therefore, the novel ERP effects reported in the present study can be considered as highly robust.

Using ERP paradigms, a number of studies have reported early brain responses to unexpected musical events. Koelsch et al. (2000, 2002b) have described an early right anterior negativity (E(R)AN) peaking around 200 ms after the unexpected chord in a harmonic chord progression and this component was shown to be generated in the inferior frontal gyrus (Maess et al., 2001). Studies with fMRI have also found activations in inferior frontal regions (mostly right lateralized) as well as in the superior temporal gyrus for violations in the harmonic regularities (Koelsch and Friederici, 2003; Koelsch et al., 2002a). Further investigations with MEG using familiar and well-memorised melodies have shown that unexpected notes elicit an early right-dominant event-related field in the auditory cortex (iMMN around 175 ms in Herholz et al., 2008; M130 in Heschl's gyrus at 130 ms in Yasui et al., 2009).

In the present study, we did not find any early ERP components which were significantly different between high-probability and low-probability notes, whereas we found novel late ERP responses. This contrast could be explained due to some clear differences between the present study and the previous ones. First, the majority of earlier investigations focused on the perception of expectations at the end of a phrase or a melody and are therefore more related to the effect of *closure*—the cohesion at the very end of a phrase or melody—than to ongoing melodic expectation, which was studied here. An exception was the recent study of Miranda and Ullman (2007) which used rule-based or memory-based violations, or a combination of both, located at any position in the musical stimulus. In this study, out-of-key violations of the rules of tonal harmony elicited an E(R)AN, whereas in-key or out-of-key memory violation triggered a late negativity, resembling the centroposterior N400. Low-probability notes in our stimuli did not constitute out-of-key violations; the unexpectedness of notes was defined purely on the basis of their conditional probability. Therefore, our results should be compared with the N400 in response to in-key memory violations in Miranda and Ullmann's study. Interestingly, the N400 typically has a bilateral centroparietal distribution. However, while the late negativity observed in our study did have a centroparietal topography, it was lateralized to the left, so the relationship between the observed late negative ERP component and N400 related processes is not clear. Furthermore, Miranda and Ullman, focused on out-of-key violation of the rules of tonal harmony rather than the more general melodic violations studied here, which may well have different spatiotemporal neural substrates. This latter argument was also proposed by Krumhansl (2003) who did not find activations of the inferior frontal gyrus against violations of melodic expectations. Finally, the melodies in the studies of Herholz et al. (2008) and Yasui et al. (2009) were either familiar or memorized, and may have involved regions responsible for memory in music. Besides, the reported iMMN was only found in musicians and not in non-musicians, such as the participants in our electrophysiological experiment.

Oscillations and phase coupling in beta band

The ERP analysis was complemented with investigation of oscillatory activity and inter-regional phase locking. Across different frequency bands, the most robust differences between low-probability and high-probability melodic events were found in the beta band. Processing low-probability notes was associated with a strong increase in beta oscillations around 500 ms over

centroparietal regions and in phase locking around the same time between electrodes located over centroparietal and occipital regions.

The convergence of results of the beta band spectral power and phase coupling strongly suggests that local neural populations showing increases in beta oscillations might be interacting by means of long-range beta phase coupling to form a large-scale network over centroparietal regions. Such long-range synchrony between spatially distant brain regions has long been postulated to mediate the formation of large-scale neural networks required for general cognitive integration (Bressler and Kelso, 2001; Fries, 2005; Varela et al., 2001). Therefore, the synchronized large-scale network of local beta oscillations reported here might have been engaged in processing the violation of pitch expectations.

To the best of our knowledge, this is the first study reporting beta oscillatory activity over the centroparietal regions during processing melodic expectations. Most previous evidence for bursts of beta oscillations is limited to motor tasks (Kuhn et al., 2004; Muller et al., 2003; Pfurtscheller et al., 2005; Pfurtscheller et al., 1997). Some new studies also speak for a role of beta oscillations in attention and cognition (Basile, 2007; Kukleta et al., 2009), although other findings rather suggest a link between reduction of beta oscillations and attention (Dalal et al., 2009; Sheth et al., 2009). In brief, increases in beta oscillations have been reported predominantly after movement execution (Muller et al., 2003; Pfurtscheller et al., 1997) and movement imagination (Kuhn et al., 2004; Muller-Putz et al., 2007; Pfurtscheller et al., 2005). Consequently, there is evidence that cortical deactivation of the motor cortex is coincident with increases in beta oscillations (Pfurtscheller et al., 1997; Salmelin et al., 1995). In the auditory domain, broad gamma band (20–60 Hz) oscillations—which also includes beta band—increase at the time of the omitted accented tone in an isochronous sequence, thereby suggesting a role in metrical interpretation (Snyder and Large, 2005). In the same line, Iversen et al. (2009) reported early beta oscillations in association with endogenous metrical interpretation. Gamma and beta rhythms together also serve as a marker of stimulus novelty (Haenschel et al., 2000; Kisley and Cornwell, 2006). But in our study, the beta band shows distinctly different results from the gamma band, and therefore, we suggest that beta and gamma bands oscillations do not necessarily represent similar dynamical properties (Kopell et al., 2000). Interestingly, MEG studies have recently linked beta oscillations to action–perception interactions. First, Iversen et al. (2009) proposed that the patterns of beta oscillations obtained were an indication that motor–auditory interactions play a role in the perceptual organization of rhythmic sound. Second, Fujioka et al. (2009) suggested that beta band activity in the auditory cortex could mediate the signalling of timing cues to facilitate motor preparatory processes for sound synchronization, such as in dancing. Given the suggested role of beta band in linking distant brain areas (Schnitzler and Gross, 2005), we postulate that the synchronized beta oscillations, as reported here, were crucial to integrate centroparietal areas into a large-scale network during processing of the low-probability note probes since the degree of integration is supposed to be higher for low-probability notes as compared to high-probability ones. Further, in the light of the works of Iversen et al. (2009) and Fujioka et al. (2009), we propose that bursts of beta oscillations during processing low-probability notes might reflect an increased auditory–motor interaction. This possible link will be focus of future investigations.

In conclusion, our results firmly establish that perceptual expectations in melody can be modelled in probabilistic terms as the output of a process of unsupervised statistical learning and, for the first time, show that melodic events associated with high and low information content are characterised by distinct patterns of neural activity. Our results further suggest that perception of melodic expectation relies on neural correlates similar to those related to auditory–motor interaction, and as such, provides new evidence about the implicit link

between perception and action in music perception (Molnar-Szakacs and Overy, 2006).

Acknowledgments

The work was partially supported by JST.ERATO project (J.B.), Goldsmiths RKTC project (J.B.), and by EPSRC research grant GR/S82220 (M.T.P., G.A.W.). We thank Rob Davis for technical help with EEG, and Michael Gale for the musical excerpts. Author contributions: M.T.P. invented the computational model; M.T.P., G.A.W., and J.B. designed the experiments; S.K. performed the experiments under the supervision of J.B. and M.T.P.; M.H.R., M.T.P., and J.B. analyzed the data; M.T.P., M.H.R., and J.B. wrote the manuscript; and all authors contributed to refining the interpretations and the final edition.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.12.019.

References

- Aarden, B., 2003. Dynamic Melodic Expectancy. Ohio State University, Columbus, OH.
- Basile, L.F., 2007. Complex, multifocal, individual-specific attention-related cortical functional circuits. *Biol. Res.* 40, 451–470.
- Bell, T.C., Cleary, J.G., Witten, I.H., 1990. Text Compression. Prentice Hall, Englewood Cliffs, N.J.
- Besson, M., Faita, F., 1995. Event-related potential (ERP) study of musical expectancy—comparison of musicians with nonmusicians. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1278–1296.
- Bharucha, J.J., Stoeckig, K., 1986. Reaction time and musical expectancy: priming of chords. *J. Exp. Psychol. Hum. Percept. Perform.* 12, 403–410.
- Bhattacharya, J., Petsche, H., Pereda, E., 2001. Long-range synchrony in the gamma band: role in music perception. *J. Neurosci.* 21, 6329–6337.
- Bressler, S.L., Kelso, J.A.S., 2001. Cortical coordination dynamics and cognition. *Trends Cogn. Sci.* 5, 26–36.
- Bunton, S., 1997. Semantically motivated improvements for PPM variants. *Comput. J.* 40, 76–93.
- Castellanos, N.P., Makarov, V.A., 2006. Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis. *J. Neurosci. Methods* 158, 300–312.
- Clarke, E.F., 2005. Ways of Listening: An Ecological Approach to the Perception of Musical Meaning. Oxford Univ. Press, Oxford, New York.
- Conklin, D., Witten, I.H., 1995. Multiple viewpoint systems for music prediction. *J. New Music Res.* 24, 51–73.
- Dalal, S.S., Baillet, S., Adam, C., Ducorps, A., Schwartz, D., Jerbi, K., Bertrand, O., Garnero, L., Martinerie, J., Lachaux, J.P., 2009. Simultaneous MEG and intracranial EEG recordings during attentive reading. *NeuroImage* 45, 1289–1304.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.
- Dowling, W.J., 1989. Simplicity and complexity in music and cognition. *Contemp. Music Rev.* 4, 247–253.
- Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480.
- Friston, K., 2005. A theory of cortical responses. *Philos. Trans. R. Soc. B-Biol. Sci.* 360, 815–836.
- Fujioka, T., Trainor, L.J., Large, E.W., Ross, B., 2009. Beta and gamma rhythms in human auditory cortex during musical beat processing. *Ann. N. Y. Acad. Sci.* 1169, 89–92.
- Good, P.I., 2005. Permutation, Parametric and Bootstrap Tests of Hypotheses, 3rd ed. Springer, New York.
- Haenschel, C., Baldeweg, T., Croft, R.J., Whittington, M., Gruzelier, J., 2000. Gamma and beta frequency oscillations in response to novel auditory stimuli: a comparison of human electroencephalogram (EEG) data with in vitro models. *Proc. Natl. Acad. Sci. U. S. A.* 97, 7645–7650.
- Hale, J., 2006. Uncertainty about the rest of the sentence. *Cogn. Sci.* 30, 643–672.
- Herholz, S.C., Lappe, C., Knief, A., Pantev, C., 2008. Neural basis of music imagery and the effect of musical expertise. *Eur. J. Neurosci.* 28, 2352–2360.
- Huron, D., 2006. Sweet anticipation: music and the psychology of expectation. MIT Press, Cambridge, MA.
- Hurtado, J.M., Rubchinsky, L.L., Sigvardt, K.A., 2004. Statistical method for detection of phase-locking episodes in neural oscillations. *J. Neurophysiol.* 91, 1883–1898.
- Hyvarinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. *Neural Netw.* 13, 411–430.
- Itti, L., Baldi, P., 2006. Bayesian surprise attracts human attention. In: Weiss, Y., Scholkopf, B., Platt, J. (Eds.), *Advances in Neural Information Processing Systems*. MIT Press, Cambridge, MA, pp. 547–554.
- Iversen, J.R., Repp, B.H., Patel, A.D., 2009. Top-down control of rhythm perception modulates early auditory responses. *Ann. N. Y. Acad. Sci.* 1169, 58–73.

- Jasper, H.H., 1958. Report of the committee on methods of clinical examination in electroencephalography. *Electroencephal. Clin. Neurophys.* 10, 371–375.
- Kisley, M.A., Cornwell, Z.M., 2006. Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clin. Neurophysiol.* 117, 2549–2563.
- Koelsch, S., Friederici, A.D., 2003. Toward the neural basis of processing structure in music—comparative results of different neurophysiological investigation methods. *Neurosci. Music* 999, 15–28.
- Koelsch, S., Gunter, T., Friederici, A.D., Schroger, E., 2000. Brain indices of music processing: “nonmusicians” are musical. *J. Cogn. Neurosci.* 12, 520–541.
- Koelsch, S., Gunter, T.C., von Cramon, D.Y., Zysset, S., Lohmann, G., Friederici, A.D., 2002a. Bach speaks: a cortical “language–network” serves the processing of music. *NeuroImage* 17, 956–966.
- Koelsch, S., Schroger, E., Gunter, T.C., 2002b. Music matters: preattentive musicality of the human brain. *Psychophysiology* 39, 38–48.
- Koelsch, S., Kilches, S., Steinbeis, N., Schelinski, S., 2008. Effects of unexpected chords and of performer's expression on brain responses and electrodermal activity. *PLoS ONE* e2631, 3.
- Kopell, N., Ermentrout, G.B., Whittington, M.A., Traub, R.D., 2000. Gamma rhythms and beta rhythms have different synchronization properties. *Proc. Natl. Acad. Sci. U. S. A.* 97, 1867–1872.
- Krumhansl, C.L., 1990. *Cognitive Foundations of Musical Pitch*. Oxford Univ. Press, New York.
- Krumhansl, C.L., 2003. Experimental strategies for understanding the role of experience in music cognition. In: Avanzini, G., Faienza, C., Minciachchi, D., Lopez, L.M.M. (Eds.), *The Neuroscience of Music*. The New York Academy of Sciences, New York, pp. 414–428.
- Krumhansl, C.L., Kessler, E.J., 1982. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* 89, 334–368.
- Kuhn, A.A., Williams, D., Kupsch, A., Limousin, P., Hariz, M., Schneider, G.H., Yarrow, K., Brown, P., 2004. Event-related beta desynchronization in human subthalamic nucleus correlates with motor performance. *Brain* 127, 735–746.
- Kukleta, M., Bob, P., Brazdil, M., Roman, R., Rektor, I., 2009. Beta-2 band synchronization during a visual oddball task. *Physiol. Meas.* 58, 725–732.
- Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208.
- Lagerlund, T.D., Sharbrough, F.W., Busacker, N.E., Cicora, K.M., 1995. Interelectrode coherences from nearest-neighbor and spherical harmonic expansion computation of Laplacian of scalp potential. *Electroencephalogr. Clin. Neurophysiol.* 95, 178–188.
- Leino, S., Brattico, E., Tervaniemi, M., Vuust, P., 2007. Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Res.* 1142, 169–177.
- Lerdahl, F., Jackendoff, R., 1983. *A Generative Theory of Tonal Music*. MIT Press, Cambridge Mass.
- Levy, R., 2008. Expectation-based syntactic comprehension. *Cognition* 106, 1126–1177.
- Loui, P., Grent-t-Jong, T., Torpey, D., Woldorff, M., 2005. Effects of attention on the neural processing of harmonic syntax in Western music. *Brain Res. Cogn. Brain Res.* 25, 678–687.
- MacKay, D.J.C., 2003. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, U.K. ; New York.
- Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2001. Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* 4, 540–545.
- Manning, C.D., Schütze, H., 1999. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA.
- Meyer, L.B., 1956. *Emotion and Meaning in Music*. The University of Chicago Press, Chicago.
- Miranda, R.A., Ullman, M.T., 2007. Double dissociation between rules and memory in music: an event-related potential study. *NeuroImage* 38, 331–345.
- Molnar-Szakacs, I., Overy, K., 2006. Music and mirror neurons: from motion to 'emotion. *Soc. Cogn. Affect. Neurosci.* 1, 235–241.
- Muller, G.R., Neuper, C., Rupp, R., Keinrath, C., Gerner, H.J., Pfurtscheller, G., 2003. Event-related beta EEG changes during wrist movements induced by functional electrical stimulation of forearm muscles in man. *Neurosci. Lett.* 340, 143–147.
- Muller-Putz, G.R., Zimmermann, D., Graimann, B., Nestinger, K., Korisek, G., Pfurtscheller, G., 2007. Event-related beta EEG-changes during passive and attempted foot movements in paraplegic patients. *Brain Res.* 1137, 84–91.
- Narmour, E., 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication–Realization Model*. University of Chicago Press, Chicago.
- Nicholson, S., Knight, G.H., Dykes Bower, J. (Eds.), 1950. *Ancient and Modern Revised*. William Clowes and Sons, Suffolk, UK.
- Nittono, H., Bito, T., Hayashi, M., Sakata, S., Hori, T., 2000. Event-related potentials elicited by wrong terminal notes: effects of temporal disruption. *Biol. Psychol.* 52, 1–16.
- Nunez, P.L., Srinivasan, R., Westdorp, A.F., Wijesinghe, R.S., Tucker, D.M., Silberstein, R.B., Cadusch, P.J., 1997. EEG coherency: 1. Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalogr. Clin. Neurophysiol.* 103, 499–515.
- Overath, T., Cusack, R., Kumar, S., Von Kriegstein, K., Warren, J.D., Grube, M., Carlyon, R.P., Griffiths, T.D., 2007. An information theoretic characterisation of auditory encoding. *Plos. Biol.* 5, 2723–2732.
- Paller, K.A., McCarthy, G., Wood, C.C., 1992. Event-related potentials elicited by deviant endings to melodies. *Psychophysiology* 29, 202–206.
- Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J., 1998. Processing syntactic relations in language and music: an event-related potential study. *J. Cogn. Neurosci.* 10, 717–733.
- Pearce, M.T., Wiggins, G.A., 2004. Improved methods for statistical modelling of monophonic music. *J. New Music Res.* 33, 367–385.
- Pearce, M.T., Wiggins, G.A., 2006. Expectation in melody: the influence of context and learning. *Music Percept.* 23, 377–405.
- Pearce, M.T., Conklin, D., Wiggins, G.A., 2005. Methods for combining statistical models of music. In: Will, U.K. (Ed.), *Computer Music Modelling and Retrieval*. Springer-Verlag, Heidelberg, Germany, pp. 295–312.
- Pereda, E., Quiroga, R.Q., Bhattacharya, J., 2005. Nonlinear multivariate analysis of neurophysiological signals. *Prog. Neurobiol.* 77, 1–37.
- Peretz, I., 2001. Brain specialization for music. New evidence from congenital amusia. *Ann. N. Y. Acad. Sci.* 930, 153–165.
- Peretz, I., 2002. Brain specialization for music. *Neuroscientist* 8, 372–380.
- Peretz, I., Hebert, S., 2000. Toward a biological account of music experience. *Brain Cogn.* 42, 131–134.
- Pfurtscheller, G., Stancak, A., Edlinger, G., 1997. On the existence of different types of central beta rhythms below 30 Hz. *Electroencephalogr. Clin. Neurophysiol.* 102, 316–325.
- Pfurtscheller, G., Neuper, C., Brunner, C., da Silva, F.L., 2005. Beta rebound after different types of motor imagery in man. *Neurosci. Lett.* 378, 156–159.
- Rodriguez, E., George, N., Lachaux, J.P., Martinerie, J., Renault, B., Varela, F.J., 1999. Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–433.
- Salmelin, R., Hamalainen, M., Kajola, M., Hari, R., 1995. Functional segregation of movement-related rhythmic activity in the human brain. *NeuroImage* 2, 237–243.
- Schellenberg, E.G., 1996. Expectancy in melody: tests of the implication–realization model. *Cognition* 58, 75–125.
- Schellenberg, E.G., 1997. Simplifying the implication–realization model of melodic expectancy. *Music Percept.* 14, 295–318.
- Schellenberg, E.G., Adachi, M., Purdy, K.T., McKinnon, M.C., 2002. Expectancy in melody: tests of children and adults. *J. Exp. Psychol. Gen.* 131, 511–537.
- Schnitzler, A., Gross, J., 2005. Normal and pathological oscillatory communication in the brain. *Nat. Rev. Neurosci.* 6, 285–296.
- Schon, D., Besson, M., 2005. Visually induced auditory expectancy in music reading: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 17, 694–705.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 623–656.
- Sheth, B.R., Sandkuhler, S., Bhattacharya, J., 2009. Posterior beta and anterior gamma oscillations predict cognitive insight. *J. Cogn. Neurosci.* 21, 1269–1279.
- Snyder, J.S., Large, E.W., 2005. gamma-band Activity reflects the metric structure of rhythmic tone sequences. *Brain Res. Cogn. Brain Res.* 24, 117–126.
- Stam, C.J., Nolte, G., Daffertshofer, A., 2007. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum. Brain Mapp.* 28, 1178–1193.
- Steiger, J.H., 1980. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* 87, 245–251.
- Steinbeis, N., Koelsch, S., Sloboda, J.A., 2006. The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *J. Cogn. Neurosci.* 18, 1380–1393.
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1997. Oscillatory gamma-band (30–70 Hz) activity induced by a visual search task in humans. *J. Neurosci.* 17, 722–734.
- Tass, P., Rosenblum, M.G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A., Freund, H.J., 1998. Detection of n : m phase locking from noisy data: application to magnetoencephalography. *Phys. Rev. Lett.* 81, 3291–3294.
- Temperley, D., 2001. *The cognition of basic musical structures*. MIT Press, Cambridge, Mass.
- Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and large-scale integration. *Nat. Rev., Neurosci.* 2, 229–239.
- Verleger, R., 1990. P3-evoking wrong notes: unexpected, awaited, or arousing? *Int. J. Neurosci.* 55, 171–179.
- Yasui, T., Kaga, K., Sakai, K.L., 2009. Language and music: differential hemispheric dominance in detecting unexpected errors in the lyrics and melody of memorized songs. *Hum. Brain Mapp.* 30, 588–601.