# Evidence for Pitch Chroma Mapping in Human Auditory Cortex

Paul M. Briley[1,2], Charlotte Breakey[3] and Katrin Krumbholz[1]

[1]MRC Institute of Hearing Research, Nottingham, UK, [2]Department of Psychology, University of York, York, UK, and [3]School of Biomedical Sciences, University of Nottingham, Nottingham, UK

Address correspondence to Paul M. Briley, Department of Psychology, University of York, York YO10 5DD, UK. Email: paul.briley@york.ac.uk.

Some areas in auditory cortex respond preferentially to sounds that elicit pitch, such as musical sounds or voiced speech. This study used human electroencephalography (EEG) with an adaptation paradigm to investigate how pitch is represented within these areas and, in particular, whether the representation reflects the physical or perceptual dimensions of pitch. Physically, pitch corresponds to a single monotonic dimension: the repetition rate of the stimulus waveform. Perceptually, however, pitch has to be described with 2 dimensions, a monotonic, "pitch height," and a cyclical, "pitch chroma," dimension, to account for the similarity of the cycle of notes (c, d, e, etc.) across different octaves. The EEG adaptation effect mirrored the cyclicality of the pitch chroma dimension, suggesting that auditory cortex contains a representation of pitch chroma. Source analysis indicated that the centroid of this pitch chroma representation lies somewhat anterior and lateral to primary auditory cortex.

**Keywords:** electroencephalography, Heschl's gyrus, musical pitch, octave similarity, stimulus-specific adaptation

## Introduction

Pitch is one of the most important perceptual features of sound. It conveys prosody and speaker identity in speech (Smith and Patterson 2005) and melody in music (Scherer 1995), and it is one of the most important cues for segregating sounds from different sources in the environment (Darwin 1997; Carlyon 2004). Most tonal sounds have temporally periodic pressure waveforms, and their pitch is determined by the waveform repetition rate, $R$ ($R$ is equal to the reciprocal of the repetition period, $P$; Fig. 1A). Thus, physically, pitch corresponds to a single, monotonic dimension ranging from low to high. Perceptually, however, pitch has 2 dimensions: a monotonic, "pitch height," dimension, reflecting the octave, within which a given note resides, and a cyclical, "pitch chroma," dimension, representing the cycle of notes within each octave (Dowling 1999; Fig. 1B). This is why music psychologists represent pitch as a helix, with the linear dimension of the helix representing pitch height and the circular dimension representing pitch chroma (Ueda and Ohgushi 1987; Fig. 1C). The distance between each 2 points within the helix reflects the perceptual similarity between the corresponding notes: vertically aligned notes have the same pitch chroma (i.e. differ by an octave; e.g. $C_4$ and $C_5$ in Fig. 1C) and are thus perceived as more similar than notes on opposite sides of the helix, which have the maximum possible chroma difference (i.e. a half-octave, or "tritone"; e.g. $C_5$ and $F_4^{\#}$).

Neuroimaging studies have suggested that anterolateral Heschl's gyrus responds more strongly to sounds with salient pitch than to sounds with no or weak pitch (Gutschalk et al. 2002; Patterson et al. 2002; Krumbholz et al. 2003; Penagos et al. 2004; Puschmann et al. 2010). Results from single-unit recordings suggest that the monkey homologue of this area contains neurons that are selective for pitch (Bendor and Wang 2005, 2010). Owing to the limitations in spatial resolution of noninvasive neuro-recording techniques, it would be difficult to measure cortical selectivity for pitch in humans using conventional stimulation paradigms. However, it has been suggested that paradigms based on adaptation might offer a means to overcome these limitations. The idea is that presentation of an "adapter" stimulus (A) produces a temporary reduction in the sensitivity of neurons responsive to that stimulus. The neural response to a subsequent "probe" stimulus (A–P) would then be assumed to reflect the representational similarity of the adapter and probe: a strongly adapted probe response would indicate that the adapter and probe are represented by the same, or similar, groups of neurons (Fig. 2A), whereas a weakly adapted probe response would indicate recruitment of mostly new, or unadapted, neurons by the probe, and thus representation by different groups of neurons (Fig. 2B). This idea has been widely used to probe sensory representations in human cortex with psychophysics (e.g. Blakemore and Campbell 1969; Kay and Matthews 1972), noninvasive electrophysiological recording techniques (electroencephalography [EEG] and magnetoencephalography [MEG]; e.g. Butler 1968; Näätänen et al. 1988) and, more recently, functional magnetic resonance imaging (fMRI; Grill-Spector and Malach 2001; see Grill-Spector et al. 2006, for review).

This study used this adaptation approach to probe the neural representation of pitch in human auditory cortex. The aim was to investigate whether adaptation of auditory cortical responses is selective for pitch and, if so, whether the selectivity reflects the perceptual similarity between notes of the same chroma. Adaptation was measured with EEG. The adapter and probe were complex tonal sounds similar to the sounds produced by most musical instruments or the voiced portions of speech. They differed from each other in terms of repetition rate, $R$, and thus pitch. If pitch is represented in terms of its physical dimension (i.e. repetition rate), the probe response should increase monotonically with increasing pitch separation between the adapter and probe. However, if pitch is represented in terms of its perceptual dimensions (pitch height and pitch chroma), the function relating the probe response size to the pitch separation, henceforth referred to as "adaptation function," should be nonmonotonic, with a dip at octave pitch separations. For comparison, we also included a condition in which the adapter and probe were sinusoids (or "pure tones") differing in frequency. Previous neurophysiological studies have assumed that the cortical coding of pitch involves dedicated "pitch neurons" that code pitch invariant of the stimulus spectral composition (Schwarz and
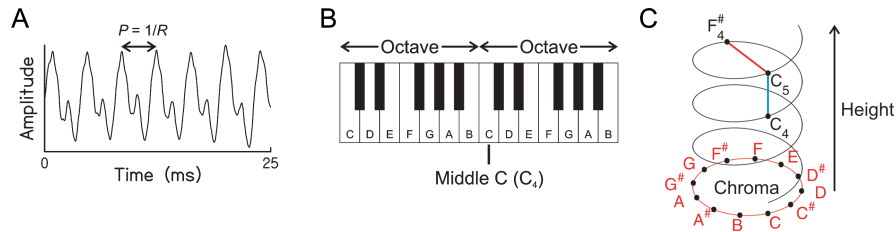
**Figure 1.** Physical and perceptual dimensions of pitch. (*A*) Pressure waveform of middle *C* played on the piano. The waveform repetition rate, *R*, is the reciprocal of the waveform period, *P*. (*B*) Schematic piano keyboard (2 octaves shown), indicating the position of middle *C*. (*C*) The pitch helix, consisting of a linear, "pitch height," dimension and a cyclical, "pitch chroma", dimension. Two notes separated by an octave (blue line) have a lesser distance than 2 notes separated by a half-octave (red line).
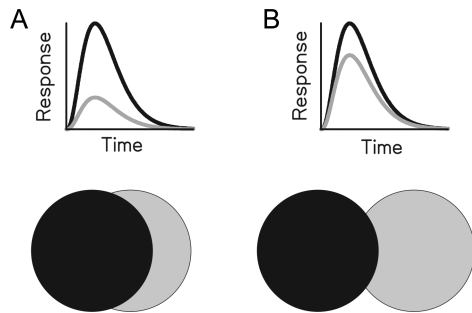


**Figure 2.** Schematic illustrating the rationale of the adaptation paradigm. (*A*) A strongly adapted probe response is taken to indicate that the adapter (black) and probe (gray) are represented by the same, or similar, groups of neurons. (*B*) A weakly adapted probe response is taken to indicate representation by different groups of neurons.

Tomlinson 1990; Fishman et al. 1998; Steinschneider et al. 1998; Bendor and Wang 2005, 2010; but see Schnupp and Bizley 2010, for a recent critique of this idea). This idea has arisen as a result of the fact that sounds with different spectral compositions, and thus different "timbre," can still elicit the same pitch. Examples include the different vowels in speech or the sounds produced by different musical instruments. Dedicated pitch neurons would be expected to be similarly activated by pure tones as by complex tones with the same pitch. Under this assumption, the pure-tone condition would be expected to yield a similar pattern of results as the complex-tone condition.

## Materials and Methods

### Stimuli

Each trial consisted of an adapter stimulus followed immediately (in order to maximize the adaptation effect) by a short (250-ms) probe stimulus (Fig. 3*A*). The stimuli were gated on and off with 10-ms quarter-cosine ramps to avoid audible clicks. At their transition, the gates were crossfaded so that the intensity envelope of the composite stimulus remained flat. The adapter was much longer than the probe (1500 ms) to allow the response to the adapter onset ("OnR" in Fig. 3*B*) to subside before the probe onset. The adapter and probe were either pure tones or a quasiperiodic noise, referred to as "iterated rippled noise," or IRN (Yost et al. 1996). All stimuli were presented binaurally at an overall level of 70 dB SPL. The silent gap between trials was 1500 ms.

Apart from being only partially periodic, IRN is similar to the harmonic complex tones used in many previous studies of pitch processing. IRN was generated using the "add-original" procedure described by Yost et al. (1996). This involves mixing a sample of random (Gaussian) noise with a copy of the same noise sample, delayed by the "quasiperiod," *P*, and then iterating the process (the current study used 16 iterations). *P* is equivalent to the period in harmonic tones.
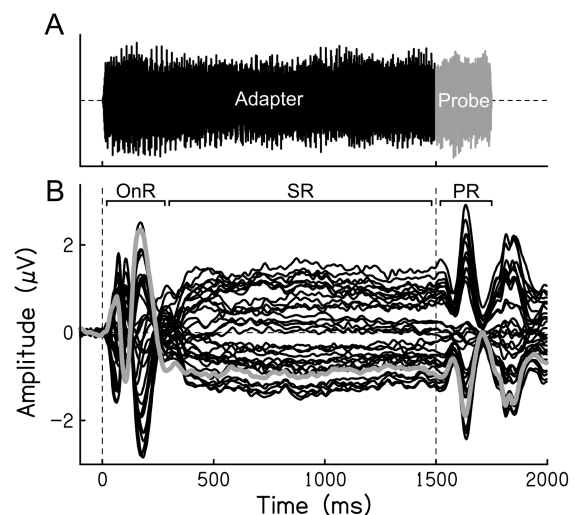


**Figure 3.** Stimulus design and average response. (*A*) Example stimulus waveform. The adapter is plotted in black, and the probe in gray. (*B*) Average response across all participants and pitch separations. Each black line represents a recording channel and the response from the vertex channel (Cz) is plotted in gray. OnR, Onset Response; SR, Sustained Response; PR, Probe Response.
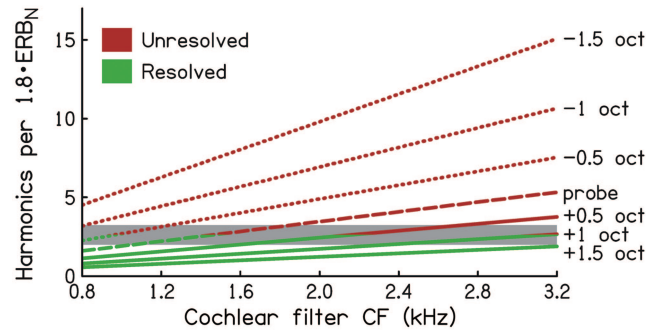
The procedure imparts a degree of periodicity to the noise waveform, which gives rise to a pitch at the reciprocal of the quasiperiod, *R* (henceforth referred to as "repetition rate"). IRN has been shown to activate similar brain areas as harmonic tones (Penagos et al. 2004) but produce a stronger pitch-related response (Barker et al. 2012). Barker et al. raised the concern that the stronger response to IRN might be related to longer-term spectro-temporal modulations that are present in IRN but not in harmonic tones. However, these modulations could not explain transient electrophysiological (EEG and MEG) responses to IRN such as those measured in this study or the study by Krumbholz et al. (2003), because these responses set in only a few tens of milliseconds after the stimulus onset (see Fig. 6), at which point the longer-term modulations have not yet unfolded. This was confirmed by Steinmann and Gutschalk (2012) using MEG; they showed that both the transient and sustained MEG responses to IRN are unaffected by the IRN modulations. The adapter and probe were generated afresh, using new noise samples, for each trial. The repetition rate (IRN) or frequency (pure tones) of the probe stimulus had a nominal value of 125 or 500 Hz, respectively. These values were chosen to ensure that the repetition rates of the IRN stimuli were within the range that is relevant for music and speech, and that the frequencies of the pure-tone stimuli were within the hearing range. The exact value of the probe repetition rate or frequency was varied from trial to trial within a one-third-octave range around the nominal value to avoid across-trial adaptation to the probe. The repetition rate or frequency of the adapters was varied relative to that of the probe to vary the pitch separation between the adapter and probe. In Experiment 1, pitch separations of 0.5, 1, and 1.5 octaves were used. The
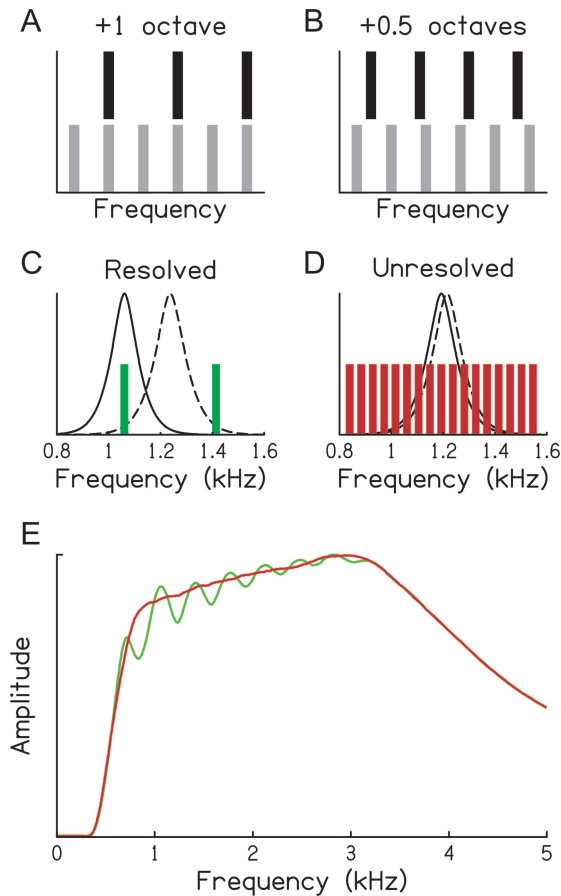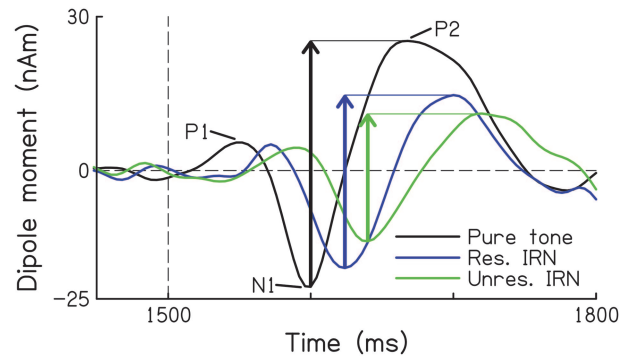
**Figure 4.** Spectral properties of IRN stimuli. (*A* and *B*) The harmonics of the adapter (black bars) and probe (gray bars) have greater overlap when the adapter and probe are separated by an octave (*A*) than a half-octave (*B*). (*C*–*E*) Resolved harmonics (*C*) stimulate separate cochlear filters and, as a result, produce peaks in the distribution of activity across the tonotopic array (*E*, green line). Multiple unresolved harmonics (*D*) fall into each cochlear filter, producing a uniform activity distribution (*E*, red line).



**Figure 5.** Degree of spectral resolvability of IRN stimuli used in Experiment 1. The ordinate shows the number of harmonics falling into one cochlear filter as a function of the filter center frequency (abscissa). The dashed line shows the probe stimulus. The dotted and solid lines show the unresolved and resolved adapters, respectively; the parameter is the pitch separation from the probe (see labels on the right). The gray horizontal bar marks the limit of harmonic resolvability according to the criterion of Shackleton and Carlyon (1994; 2–3.25 harmonic per filter). Above this limit, harmonics are unresolved (red line segments) and below, harmonics are resolved (green line segments).



**Figure 6.** Average probe response to each stimulus condition in Experiment 1. The arrows illustrate how the response sizes and latencies were measured.

0.5- and 1.5-octave conditions had the same pitch chroma separation (a half-octave, or tritone, modulo one octave), and their average pitch height separation matched that of the 1-octave condition.

IRN contains spectral peaks at frequencies corresponding to integer multiples ("harmonics") of the stimulus repetition rate. This creates potential for confound, because 2 notes separated by an octave have greater spectral overlap (they share every other harmonic, see Fig. 4*A*) than notes separated by a half-octave (their harmonics are nonoverlapping; Fig. 4*B*). A smaller probe response for the 1-octave than half-octave conditions might thus arise as a result of stronger adaptation of frequency-selective neurons, rather than selectivity to pitch chroma. However, this confound only applies when the harmonics are resolved by the cochlear frequency filters (i.e. the spacing between adjacent harmonics is greater than the widths of the filter tuning curves; Fig. 4*C*). Resolved harmonics produce peaks in the pattern of activity across the tonotopic map (green line in Fig. 4*E*), whereas unresolved harmonics (i.e. each cochlear filter responds to multiple harmonics; Fig. 4*D*) produce a uniform activity distribution (red line in Fig. 4*E*). The frequency tuning width of the cochlear filters increases roughly proportionally with the filter frequency (Glasberg and Moore 1990). This means that, for a given harmonic sound, only harmonics up to about the 10th are resolved (Shackleton and Carlyon 1994). It also means that, for a given frequency band, harmonic sounds with repetition rates below about 1/10th of the lower edge of the band are unresolved across the entire band, and sounds with repetition rates above 1/10th of the lower edge are at least partially resolved. To investigate the effect of harmonic resolvability, we used both resolved and unresolved IRN adapters.

Any difference in probe response size between the 1-octave and half-octave conditions for the unresolved IRN adapters would have to be assumed to reflect the properties of pitch-selective neurons. The IRN stimuli were bandpass-filtered between 800 Hz and 3.2 kHz using an eighth-order Butterworth IIR filter (which yields a −24-dB/oct filter roll-off). This meant that the IRN probe was resolved within about the lower third of the stimulus passband (see dashed line in Fig. 5). The repetition rates of the adapters were either above (solid lines in Fig. 5) or below (dotted lines) the probe repetition rate. The adapters with the higher rates (+0.5, +1, +1.5 octaves) were resolved within at least about half of the passband, and the adapters with the lower rates (−0.5, −1, −1.5 octaves) were unresolved over practically the entire band.

In the pure-tone conditions, which were used in Experiments 1 and 3 (see below), adapter frequencies both above and below the probe frequency were used for each pitch separation (0.5, 1, 1.5 octaves for Experiment 1, and 1.5 octaves for Experiment 3).

Both the IRN and pure-tone stimuli were presented in a background of masking noise, presented continuously throughout the data acquisition. The masker for the IRN stimuli was intended to prevent audible distortion products below the stimulus passband. It was lowpass-filtered at a half-octave below the lower edge of the band (i.e. 566 Hz) using an eighth-order Butterworth filter as before (−24-dB/oct filter roll-off) and presented at a level of 40 dB SPL per cochlear-filter bandwidth (as defined by Glasberg and Moore 1990). The masker for the pure-tone stimuli was intended to approximately equalize the level above detection threshold (sensation level), and

thus the loudness, of the stimuli across the different adapter frequencies used. It was presented at a level of 30 dB SPL per cochlear-filter bandwidth.

Stimuli were generated digitally at a sampling rate of 25 kHz using Matlab (The Mathworks). They were digital-to-analogue converted with a TDT System 3 (consisting of an RP2.1 real-time digital signal processor and an HB7 headphone amplifier; Tucker David Technologies) and presented through K 240 DF headphones (AKG).

### Procedure

Experiment 1 investigated whether adaptation to pitch is sensitive to pitch chroma and consisted of 2 sessions, one for the IRN stimuli and one for the pure tones. In both sessions, stimuli were presented in 4 approximately 20-min blocks. Each block contained 372 trials (62 for each pitch separation). The pitch separations were presented in random order.

Experiment 2 investigated whether the nonmonotonicity of the adaptation functions for the IRN stimuli was due to a pitch chroma or a consonance effect. It consisted of a single session with 4 blocks. Each block contained 434 trials (31 for each of the 14 pitch separations used) and lasted approximately 24 min. As in Experiment 1, the pitch separations were presented in random order.

Experiment 3 estimated the source locations of the probe responses measured in Experiments 1 and 2. The stimuli were presented in a single session consisting of 6 blocks, 2 for the unresolved IRN stimuli (−1.5-octave pitch separation between the adapter and probe), 2 for the resolved IRN stimuli (+1.5-octave pitch separation) and 2 for the pure tones (1 with a pitch separation of −1.5 octaves, and the other with a +1.5-octave pitch separation). The blocks for the resolved IRN stimuli and the pure tones each contained 250 trials and lasted approximately 14 min. To compensate for the smaller probe response size for the unresolved IRN stimuli, the blocks for this condition contained 400 trials and lasted approximately 22 min. The 6 blocks were presented in a random order.

### Data Acquisition

Auditory-evoked cortical potentials were recorded from 32 (Experiments 1 and 2) or 64 (Experiment 3) Ag/AgCl ring electrodes (Easycap, Herrsching, Germany). The 32 electrodes were placed according to the standard 10–20 arrangement (Jasper 1958). The 64 electrodes were placed according to an extended 10–20 arrangement that provided greater coverage of the lower half of the head surface ("Infracerebral" cap, Easycap). In all experiments, the recording reference was the vertex electrode (Cz) and the ground electrode was placed on the central forehead (AFz). Skin-to-electrode impedances were kept below 5 kΩ throughout the recordings. The electrode signals were amplified with BrainAmp DC EEG amplifiers (Brain Products) and bandpass-filtered online between 0.1 and 250 Hz. The signals were sampled at 500 Hz and stored for offline analysis using the Brain Vision Recorder software (Brain Products). Participants watched a subtitled movie throughout the recordings to remain alert.

### Data Analysis

The data from Experiments 1 and 2 were preprocessed using the EEGLAB toolbox (Delorme and Makeig 2004), which runs under Matlab. They were 1) lowpass-filtered at 35 Hz using a −48-dB per octave zero-phase IIR filter, 2) down-sampled to 250 Hz, 3) re-referenced to average reference, 4) segmented into 2350-ms epochs ranging from 100 ms before the start of the adapter to 500 ms after the end of the probe, and 5) baseline-corrected to the 100-ms prestimulus period. Epochs containing unusually large potentials across many electrodes (outside of ±3 SD) were rejected using EEGLAB's "joint-probability" function. This led to the rejection of an average of 15% of epochs in Experiment 1, and 17% in Experiment 2. The remaining epochs were submitted to an independent component analysis (extended infomax algorithm; Bell and Sejnowski 1995; Lee et al. 1999) for each run and each participant separately. Components representing eye blinks, lateral eye movements, and electrocardiac

activity were removed by manual inspection. Epochs were then averaged for each participant and condition. The averaged responses were converted from sensor to source space using the Brain Electrical Source Analysis software (BESA, Gräfelfing). The source model consisted of 2 equivalent current dipoles placed at the centroids of primary area TE1.0 in the left and right hemispheres (Morosan et al. 2001). A 4-shell ellipsoidal volume conductor was used as a head model. The dipole orientations were fitted to the average probe response across conditions and participants using a time window encompassing the P1, N1, and P2 deflections (0–300 ms after probe onset). The resulting source model was used as a spatial filter to create 2 source waveforms for each condition and participant, 1 for the dipole in each hemisphere. The source waveforms were averaged across hemispheres to improve the response-to-noise ratio (none of the subsequent statistical analyses yielded any interaction with hemisphere; all $P > 0.05$).

The data from Experiment 3 were preprocessed using BESA. For each participant, the data from all 6 runs were concatenated and searched for potentials resembling eye blinks or lateral eye movements. The potentials for the eye blinks and lateral eye movements were averaged separately, and their first spatial principal components were used to define the respective artifact topographies. In order to correct for ocular artifacts, the artifact topographies were incorporated into the subsequent source model (described in Results section). As for Experiments 1 and 2, the data were segmented into 2350-ms epochs from 100 ms before the adapter onset to 500 ms after the probe offset. Epochs with voltages exceeding ±120 μV were discarded. On average, 10% of epochs were removed. The remaining epochs were averaged for each condition and participant. The parameters for the source modeling were the same as for Experiments 1 and 2, apart from the dipoles being unconstrained in both orientation and location and the fitting being based on the individual rather than the grand-average responses.

The size of the probe responses was measured, in first instance, using the N1-P2 peak-to-peak difference (see arrows in Fig. 6). The N1 and P2 deflections have opposite polarities and partly overlapping time courses (Näätänen and Picton 1987; Makeig et al. 1997), and may thus partially cancel each other. Using the peak-to-peak, rather than baseline-to-peak, measure of the probe response size avoids this cancellation from affecting the data pattern. Many of the earlier studies that have measured stimulus selectivity of adaptation in the auditory-evoked cortical potentials have taken a similar approach (reviewed in Näätänen and Picton 1987). Subsequently, we also measured the sizes of the N1 and P2 peaks separately to examine whether they showed similar effects. The latency of the probe response was taken as the N1 peak latency.

In Experiment 1, the effect of pitch height was tested by comparing the probe response sizes for the 0.5- and 1.5-octave pitch separations, and the effect of pitch chroma was tested by comparing the response size for the 1-octave pitch separation with the mean response size for the 0.5- and 1.5-octave separations.

In Experiment 2, the pitch height and pitch chroma effects were assessed by fitting the adaptation functions for the IRN conditions with a combined sinusoidal and linear function of pitch separation (red dashed lines in Fig. 7H, I). According to the pitch helix model, the sinusoidal component represents the pitch chroma distance, and the linear component the pitch height distance, between the adapter and probe. The function was defined by $P(\Delta R) = a \cdot |\sin \pi \cdot \Delta R| + b \cdot \Delta R + c$, where $P$ is the probe response size, $\Delta R$ is the pitch separation in octaves, and $a$, $b$, and $c$ are free parameters; $a$ and $b$ are scaling factors for the sinusoidal and linear function components, respectively, and $c$ is a constant offset.

### Participants

Experiment 1 was conducted with a total of 15 participants (7 males, mean ± SD: 23.5 ± 4.9 years), 6 of whom only completed the IRN session (4 blocks) and 5 only completed the pure-tone session (4 blocks). Four participants completed both sessions (all 8 blocks) on different days. Twelve participants (6 male, age: 22.1 ± 5.6 years) took part in Experiment 2 (4 blocks), and 8 participants (3 male, age:
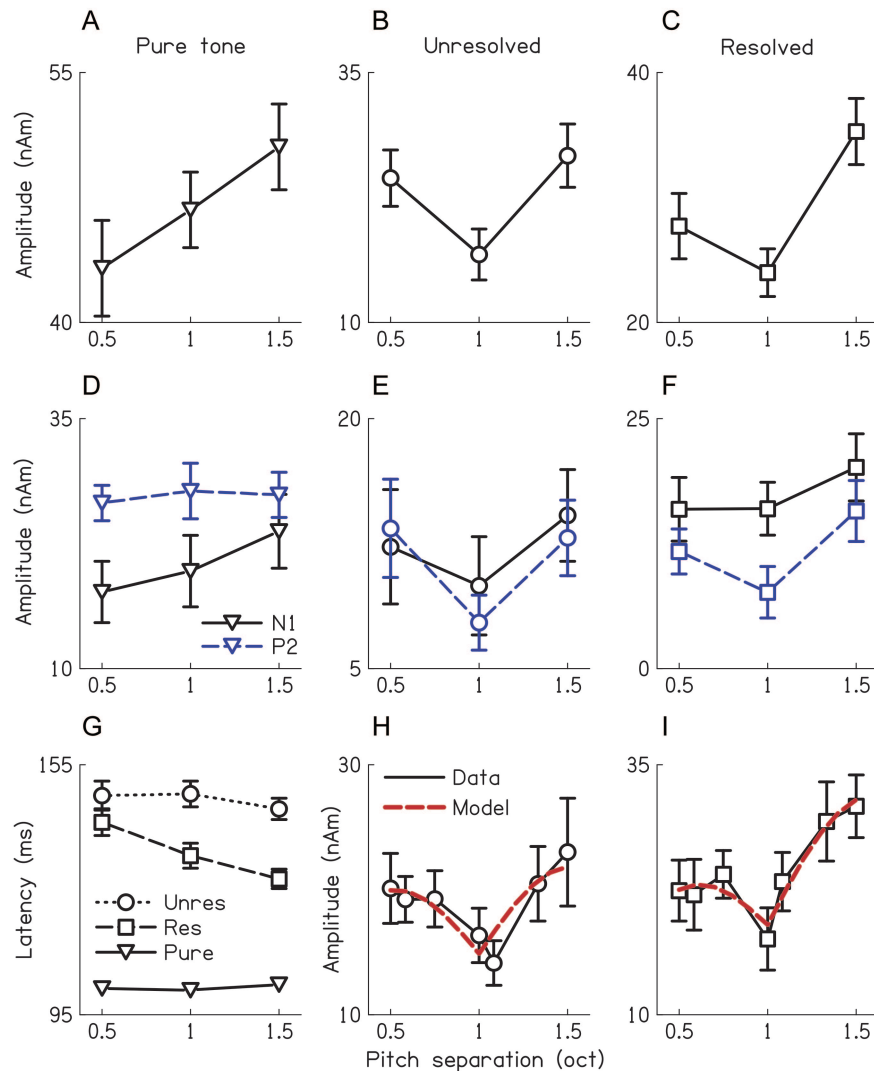
**Figure 7.** Average probe response sizes and latencies for Experiments 1 and 2. Error bars denote ± 1 standard error of the mean (SEM). (*A–C*) Probe response sizes for the pure-tone (*A*), unresolved (*B*), and resolved (*C*) IRN conditions in Experiment 1, averaged across participants and plotted as a function of the pitch separation between the adapter and probe. (*D–F*) Probe response sizes as in panels *A–C*, but broken down into contributions from the N1 (black solid lines) and P2 (blue dashed lines) deflections. (*G*) Average probe response latencies from Experiment 1 as a function of pitch separation. (*H* and *I*) Probe response sizes for the unresolved and resolved IRN conditions in Experiment 2 (black lines and symbols) with function fits representing the helical model of pitch perception (red dashed lines).

22.0 ± 2.0 years) in Experiment 3 (6 blocks). The participants in Experiment 2 and in the IRN session of Experiment 1 were nonoverlapping.

All participants had hearing thresholds of 20 dB HL or better at audiometric frequencies between 250 and 4000 Hz, and had no history of audiological or neurological disease. The participants in Experiment 3 were screened for large EEG responses using a short (100-ms) 1000-Hz tone pip, presented at a rate of 1 per 1.5 s, as test stimulus. Participants with vertex N1 amplitudes of less than 7 µV (using a linked-mastoid reference) were excluded. Participants gave written informed consent. The study procedures were approved by the Ethics Committee of the University of Nottingham School of Psychology.

## Results

### *Dependence of Adaptation on Pitch Separation*

The probe response ("PR" in Fig. 3*B*) had a similar triphasic morphology as the adapter onset response ("OnR"), with a small initial positive peak (referred to as "P1"; Näätänen and Picton 1987), a large negative peak ("N1") and another larger positive peak ("P2"). Overall, the pure-tone condition yielded the largest and earliest probe responses (black line and arrow in Fig. 6), followed by the resolved (blue) and then unresolved (green) IRN conditions.

A linear mixed model (LMM) analysis of the probe response sizes (measured as the N1–P2 peak-to-peak difference) for the pure-tone condition in Experiment 1 was conducted to test for any effects of the frequency difference direction (adapter above or below the probe) between the adapter and probe (fixed factors: frequency difference direction and adapter-probe pitch, or frequency, separation, entered as covariate; random factor: participants). Although the probe responses were larger for adapter frequencies above than below the probe frequency (main effect of frequency difference direction: $F(1,43) = 55.498$, $P < 0.001$), the frequency difference direction had no effect on the pattern of results across pitch separations (frequency difference direction by

pitch separation interaction: $F_{(1,42)} = 0.079$, $P = 0.780$). An LMM analysis of the probe response latencies also showed no significant interaction or main effect of frequency difference direction (all $P > 0.05$). Therefore, the pure-tone probe responses from Experiments 1 and 3 were averaged across frequency difference direction. The size of the responses to the pure-tone probes increased monotonically with increasing pitch separation from the adapter (main effect of pitch separation: $F_{(1,17)} = 19.548$, $P < 0.001$; Fig. 7A). The amount of increase was similar between the 0.5- and 1-octave pitch separations and between the 1- and 1.5-octave pitch separations (as shown by the normality of the residuals from the covariance analysis, confirmed with a Shapiro-Wilk $W$ test: $W = 0.964$, $P = 0.457$).

In contrast, the probe response size for the IRN conditions was related nonmonotonically to the pitch separation between the adapter and probe, with the response size for the 1-octave pitch separation being significantly smaller than the average response size for the 0.5- and 1.5-octave pitch separations (LMM analysis with fixed factors pitch chroma and spectral resolvability; main effect of chroma: $F_{(1,28)} = 29.865$, $P < 0.001$; Fig. 7B, C]. This suggests that adaptation for IRN stimuli is influenced by pitch chroma, with adaptation being stronger when the adapter and probe have similar chroma and weaker when they have dissimilar chroma. Importantly, the difference in probe response size between the 1-octave and half-octave pitch separations was similar for the resolved and unresolved conditions (chroma by resolvability interaction: $F_{(1,27)} = 0.026$, $P = 0.874$).

The adaptation functions for the IRN stimuli also showed an effect of pitch height, in that the probe responses for the 1.5-octave pitch separation were generally larger than those for the 0.5-octave separation (LMM analysis with fixed factors pitch height and spectral resolvability; main effect of height: $F_{(1,27)} = 12.131$, $P = 0.002$). The pitch height effect depended on the resolvability of the stimuli (height by resolvability interaction: $F_{(1,27)} = 6.329$, $P = 0.018$), in that it was significant for the resolved ($P < 0.001$), but not for the unresolved ($P = 0.500$), condition (compare Fig. 7B and C).

The response latencies for the pure-tone and unresolved IRN conditions were practically independent of pitch separation (Fig. 7G). The latencies for the unresolved IRN condition were much longer than those for the pure-tone condition (147 vs. 101 ms, on average). The latencies for the resolved IRN condition were intermediate (134 ms on average) and varied as a function of the pitch separation, being similar to the latencies for the unresolved IRN stimuli at the 0.5-octave pitch separation (141 ms) and approaching the pure-tone latencies at the 1.5-octave pitch separation (128 ms). An LMM analysis of the probe response latencies showed a significant main effect of stimulus condition ($F_{(2,70.959)} = 71.369$, $P < 0.001$; $P < 0.001$ for all pair-wise comparisons), and a significant interaction with pitch separation (entered as covariate; $F_{(4,67.667)} = 7.535$, $P = 0.001$).

Separate analyses of the N1 and P2 peaks showed that the pitch separation effect in the pure-tone condition was driven mainly by the N1 (main effect of pitch separation: $F_{(1,17)} = 27.623$, $P < 0.001$; Fig. 7D); the effect was nonsignificant for the P2 ($F_{(1,17)} = 0.202$, $P = 0.659$). The pitch chroma effect for the IRN conditions was found in both the N1 (main effect of chroma: $F_{(1,28)} = 5.273$, $P = 0.029$) and the P2 ($F_{(1,28)} = 20.983$, $P < 0.001$; Fig. 7E, F). As for the N1–P2 difference, the

chroma effect was independent of the adapter spectral resolvability for both the N1 (chroma by resolvability interaction: $F_{(1,27)} = 0.285$, $P = 0.598$) and the P2 ($F_{(1,27)} = 0.086$, $P = 0.771$). The pitch height effect in the IRN conditions was driven mainly by the P2 (height by resolvability interaction: $F_{(1,27)} = 5.081$, $P = 0.033$; Fig. 7E, F). The effect was nonsignificant for the N1 ($F_{(1,27)} = 0.531$, $P = 0.473$).

### Pitch-Chroma or Musical Consonance Effect?

The first experiment yielded nonmonotonic adaptation functions for the IRN stimuli, with a dip at the 1-octave pitch separation compared with the 0.5- and 1.5-octave separations (Fig. 7B, C). Experiment 2 tested the possibility that, rather than reflecting a pitch chroma effect, this nonmonotonicity arose as a result of the octave being a consonant (i.e. "pleasant"), and the half-octave a dissonant ("unpleasant"), interval (Schellenberg and Trehub 1994; McDermott et al. 2010). For that the resolved and unresolved IRN conditions were remeasured with a larger set of pitch separations (6, 7, 9, 12, 13, 16, and 18 semitones). We also used a different group of participants to test the robustness of the effect. The new set of pitch separations included perfectly consonant (perfect fifth, octave), imperfectly consonant (major sixth, major third modulo 1 octave) and dissonant intervals (tritone, minor second and tritone modulo 1 octave; Table 1). If the size of the adaptation effect is determined by the degree of consonance between the adapter and probe, the new adaptation functions should exhibit dips and peaks at consonant and dissonant intervals, respectively. This, however, was not the case (Fig. 7H, I); there was no significant difference in probe response size between the consonant and dissonant intervals (LMM analysis with fixed factors consonance and resolvability; main effect of consonance: $F_{(1,154)} = 1.448$, $P = 0.231$), and the correlation between the probe response sizes and consonance ratings from McDermott et al. (2010) was also nonsignificant ($\rho_{(14)} = -0.299$, $P = 0.298$). Instead, the probe response size was a smooth, combined function of the pitch chroma and pitch height separations between the adapter and probe. According to the helical model of pitch perception (see Fig. 1C), the perceptual distance between 2 notes is a combined sinusoidal and linear function of their pitch separation, with the sinusoidal component representing the distance in pitch chroma, and the linear component representing the distance in pitch height. This model provided an excellent fit to the current data (red dashed lines in Fig. 7H, I), explaining 92.7% of variance for the resolved, and 78.5% for the unresolved, IRN stimuli (see Materials and Methods for the model implementation). The model's sinusoidal (pitch-chroma) component was significant for both the resolved ($F$-test; $F_{(1,4)}$

**Table 1**
Adapter-probe pitch separations used in Experiment 2

| Pitch separation (semitones) | Musical interval | Consonance |
|---|---|---|
| 6 | Tritone | Dissonant |
| 7 | Perfect fifth | Perfect |
| 9 | Major sixth | Imperfect |
| 12 | Octave | Perfect |
| 13 | Minor second modulo 1 octave | Dissonant |
| 16 | Major third modulo 1 octave | Imperfect |
| 18 | Tritone modulo 1 octave | Dissonant |

The second and third columns show the musical intervals constituted by the pitch separations and their degree of consonance.

= 27.560, $P = 0.006$) and unresolved conditions ($F(1,4) = 13.895$, $P = 0.020$). In contrast, the linear (pitch-height) component was only significant for the resolved condition ($F(1,4) = 29.055$, $P = 0.006$), but nonsignificant for the unresolved condition ($F(1,4) = 1.693$, $P = 0.263$). This is consistent with the findings from Experiment 1.

### Source Locations

The third experiment sought to estimate the source locations of the probe responses measured in Experiments 1 and 2. To maximize the response-to-recording noise ratio (required for accurate source localization), only the largest pitch separation (1.5 octaves) was used, which had yielded the largest responses in the first 2 experiments (Fig. 7), and a large number of trials were collected for averaging. The set of recording locations was also extended to cover a greater proportion of the lower half of the head surface and thereby facilitate source localization of activity in the region of auditory cortex (see Materials and Methods).
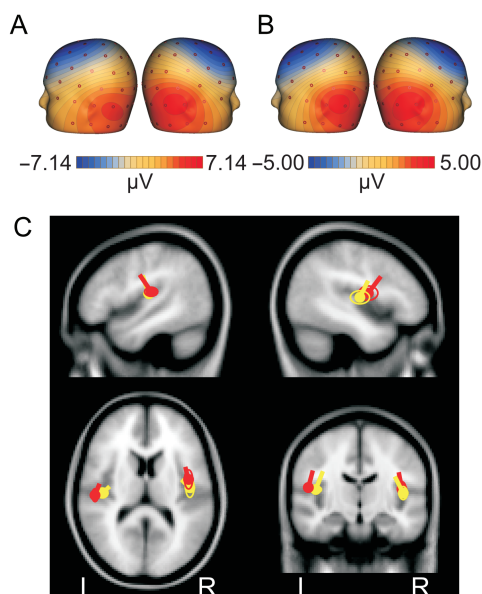
Source locations of EEG responses are derived from the responses' voltage distributions across the head surface, referred to as voltage maps. The voltage maps of all probe responses (measured over the 40-ms time window around the N1 peak; Fig. 8A, B) exhibited negative polarity around the vertex (Cz) and polarity inversion at the mastoids, indicating source locations in the general region of supratemporal auditory cortex. Consequently, the voltage map for each participant and condition was fitted with a source model consisting of 2 equivalent current dipoles, which were unconstrained in both location and orientation (Fig. 8C). Each dipole models the neural activity in a circumscribed region of cortex (in this

case, supratemporal auditory cortex in the left or right hemisphere). The dipole location reflects the centroid of the active region and the dipole orientation the direction of its net current flow (Scherg 1990). Two of the fitted dipoles were located at the boundary of the head model and were excluded from subsequent analysis. The locations and orientations of the remaining dipoles (goodness of fit ≥ 98%) were submitted to a permutation procedure (with 1000 resamples; Efron and Tibshirani 1993) to test for differences between the stimulus conditions.

The dipoles for the pure-tone condition were located on medial Heschl's gyrus (Talairach coordinates [left/right]: −41.9, −18.8, 15.8/44.2, −13.4, 13.4 mm), close to the centroid of primary area TE1.0 (Morosan et al. 2001). The largest differences in source location were observed between the pure-tone and unresolved IRN conditions. The Euclidean distance between the dipole locations for these conditions was significant in both hemispheres (left: $P = 0.024$, right: $P = 0.047$). Compared with the dipoles for the pure-tone condition (shown in yellow in Fig. 8C), the dipoles for the unresolved IRN condition (shown in red) were located 7.2 mm more lateral in the left hemisphere and 7.9 mm more anterior in the right hemisphere (Talairach coordinates [left/right]: −49.1, −21.2, 17.2/42.9, −5.5, 17.6 mm). Permutation tests showed that both differences were significant (left: $P < 0.001$, right: $P = 0.026$). Differences in dipole orientation were analyzed by treating the dipoles as vectors in the 3-dimensional unit sphere. The angle subtended by the arc between the vector endpoints is the "central angle." The central angle between the dipole orientations for the pure-tone and unresolved IRN conditions was significant in both hemispheres (left: $P < 0.001$; right: $P = 0.040$). This was due to the sagittal and transversal projections of the dipoles for the unresolved IRN condition being significantly more forward pointing than those for the pure-tone condition (sagittal [left/right]: $P = 0.014/0.016$; transversal [left/right]: $P = 0.013/0.037$). The dipole locations and orientations for the resolved IRN condition lay between those for the pure-tone and unresolved IRN conditions (Talairach coordinates [left/right]: −47.4, −21.9, 17.4/43.0, −5.3, 17.1 mm). Note that the reported Talairach coordinates are based on standard electrode positions and should thus be viewed as approximations. This does not, however, affect the observed differences between conditions.



**Figure 8.** Voltage distribution maps and dipole source models for the probe responses from Experiment 3. (A and B) Voltage maps of one representative participant for the pure-tone (A) and unresolved IRN conditions (B). (C) Dipole source models for the pure-tone (yellow) and unresolved IRN (red) conditions, averaged across participants and projected onto sections of the MNI (Montreal Neurological Institute) template brain. Ellipses show 95% confidence intervals for the dipole locations, computed using a bootstrapping procedure. The dipoles for the resolved IRN condition are not shown to avoid cluttering. The locations and orientations of the resolved IRN dipoles fell between those for the pure-tone and unresolved IRN conditions.

### Discussion

This study used an adaptation paradigm with EEG to investigate whether the pitch of complex tonal sounds, such as voiced speech, or music, is represented by its physical dimension (i.e. the waveform repetition rate) or by its perceptual dimensions (pitch height and pitch chroma) in human auditory cortex. The adaptation approach is based on the assumption that those neurons that respond most strongly to the adapter are also most adapted by it (see Grill-Spector et al. 2006, for review). According to this assumption, the amount of adaptation for a given adapter and probe should be determined by the overlap between, and thus the selectivity of, their neural representations. The most important finding of this study was that the adaptation functions for the IRN stimuli were nonmonotonic, with adaptation being stronger (i.e. the probe response being smaller) when the adapter and

probe were separated by an octave than a half-octave, or tritone. This suggests that a note and its octave share greater overlap in neural representation than a note and its half-octave. Experiment 2 ruled out the possibility that this non-monotonicity was due to the octave being a more consonant interval than the half-octave, indicating that it represents a true pitch chroma, rather than a consonance, effect. Importantly, the effect was as large for the unresolved IRN stimuli as for the resolved stimuli. As unresolved stimuli produce a uniform activity distribution across the tonotopic array, this rules out the possibility that the pitch chroma effect was due to there being greater harmonic overlap between notes with similar than dissimilar pitch chroma. These results suggest that human auditory cortex contains neurons that are selective for pitch chroma.

It is unlikely that the probe responses reflect processes involved in auditory deviance detection. The adapters and probes were presented with equal probability, and each probe was preceded by only a single adapter. Previous work suggests that such conditions are ineffective in eliciting the predictive processes that are thought to underlie the auditory deviance, or mismatch, response (e.g. Sams et al. 1983; Cowan et al. 1993; Winkler et al. 1996). Predictive processing would be expected to depend on the perceptual dissimilarity between the adapter and probe. Butler (1968) and Megela and Teyler (1979) found that adaptation in the N1–P2 amplitude is inconsistent with this expectation. They showed that a loud adapter is more effective at suppressing the response to a quiet probe than vice versa, despite the perceptual dissimilarity between the adapter and probe being the same in both cases. Results by Wacogne et al. (2011) suggest that predictive processing related to the auditory deviance response involves areas in frontal and other associative cortices. In this study, nonauditory contributions to probe responses were marginal; a principle component analysis within a time window encompassing all 3 deflections (P1, N1, and P2) of the grand-average probe response in Experiment 3 showed that a single spatial component explained over 98% of the variance in that response. This is consistent with Garrido et al.'s (2007) finding that top–down modulation of auditory-evoked cortical responses from nonauditory sources only becomes apparent after about 220 ms into the response.

In contrast to the responses to the IRN stimuli, the pure-tone responses increased monotonically with increasing pitch (or frequency) separation between the adapter and probe. This suggests that they were produced by different generators. The fact that the pure-tone responses occurred at much shorter latencies than the IRN responses suggests that they were generated at a lower processing level. The source analysis results from Experiment 3 indicated that the source of the pure-tone responses was centered on medial Heschl's gyrus, suggesting that the responses were generated in primary auditory cortex. Primary auditory cortex is known to contain a topographic representation of frequency (referred to as "tonotopic map"; Formisano et al. 2003; Talavage et al. 2004). The fact that the size of the pure-tone responses increased linearly with increasing frequency separation "in octaves" suggests that the gradient of the tonotopic map in human primary auditory cortex, like that of the cochlear tonotopic map, represents logarithmic frequency.

The source of the IRN responses was located somewhat anterior and lateral to the source of the pure-tone responses,

suggesting that it was part of a network of nonprimary areas identified as being specifically sensitive to pitch or pitch change (melody) by previous neuroimaging and neurophysiological studies (Patterson et al. 2002; Penagos et al. 2004; Gutschalk et al. 2004; Bendor and Wang 2005, 2010; Hall et al. 2006; Puschmann et al. 2010). Our results suggest that this network contains a parametric representation of pitch chroma. This is consistent with the finding by Warren et al. (2003) that a region anterior to Heschl's gyrus responded more strongly to tonal sequences that changed in pitch chroma than pitch height.

Although the resolved and unresolved IRN responses showed a similar pitch chroma effect, it is likely that the resolved responses constituted a mixture of contributions from both the pitch-sensitive nonprimary source and the frequency-selective primary source. This is suggested by the intermediate response latencies and source locations for the resolved IRN condition. The variation in response latency with pitch separation suggests that the relative proportions of the primary and nonprimary contributions to the resolved IRN responses varied as a function of pitch separation, with the primary contribution increasing with increasing pitch separation (and thus spectral dissimilarity) between the adapter and probe. The same mechanism probably also accounts for the pitch height effect observed in the resolved IRN condition (i.e. the fact that the response was larger for the 1.5- than 0.5-octave pitch separation).

The absence of any pitch chroma effect in the pure-tone condition suggests that, despite eliciting pitch, the pure tones did not evoke any notable response from the pitch-sensitive nonprimary source. This runs counter to the idea that auditory cortex contains dedicated pitch neurons that code pitch invariant of the stimulus spectral properties, or timbre. Invariant pitch neurons would have been expected to respond to IRNs and pure tones alike. Our results are consistent with the results from a seminal study by Butler (1972), who showed that a pure tone with a given pitch does not adapt the response to a complex tone with the same pitch but nonoverlapping spectral composition. The results of Butler's study and our results imply that pure tones and complex tones activate different neurons in auditory cortex. This is consistent with the finding that pure tones are an inefficient stimulus for driving nonprimary auditory neurons (Schreiner and Cynader 1984; Rauschecker et al. 1995; Wessinger et al. 2001; Hall et al. 2002) as well as the failure by previous neurophysiological studies to find neurons in primary auditory cortex that respond to the pitch of complex tones with frequency components outside of the neurons' frequency response areas (Schwarz and Tomlinson 1990; Fishman et al. 1998; Steinschneider et al. 1998). Bendor and Wang's (2005, 2010) studies represent an exception to this failure, but there is still some uncertainty as to whether their results can be attributed to distortion products, which arise as a result of the nonlinearity of cochlear processing (McAlpine 2004; Abel and Kössl 2009).

Taken together, the current and previous results suggest that in mammalian auditory cortex, pitch is corepresented together with the stimulus spectrum (or timbre), rather than being represented separately in a dedicated map. Recent studies by Nelken et al. (2008) and Bizley et al. (2009) suggest that other sound features, such as spatial location, may also be included in this representation. Their results showed that most neurons in both primary and nonprimary

auditory fields are sensitive to specific combinations of pitch, timbre, and location. This is similar to the visual cortex, where neurons represent combinations of features such as retinal location, orientation, and ocular dominance (Hubel and Wiesel 1977). It is possible that like the coding of certain visual features, such as faces or objects, pitch coding might become more specialized, and thus invariant to other features, at higher levels of processing. These levels might lie beyond the levels that generate the N1 and P2 deflections measured in this study. Alternatively, their activation might occur only under active listening conditions.

Although both the N1 and P2 showed the pitch chroma effect for the IRN stimuli, only the N1 showed the frequency separation effect for the pure tones. The N1 and P2 covary along many stimulus dimensions, which is why they have often been treated as a unitary phenomenon. However, the N1 and P2 are affected differently by attention and sleep, show different maturational time courses and have somewhat distinct topographies. This suggests that their generators are at least partially separate (see Crowley and Colrain 2004, for review). Epicortical and intracortical recordings in the rat auditory cortex suggest that the N1 reflects responses to frequency-selective thalamocortical input, whereas the P2 is generated by more widespread corticocortical connections (Barth and Di 1990; Barth et al. 1993). This may explain why only the N1, but not the P2, showed the frequency separation effect in the pure-tone condition.

A recent study by Baumann et al. (2011), which used fMRI to investigate pitch mapping in macaque monkeys, found that, at the level of the inferior colliculus (IC), pitch is mapped monotonically, with the represented pitch changing progressively from one end of the map to the other. This suggests that the IC represents the physical dimension of pitch (waveform repetition rate). Our finding that adaptation in auditory cortex shows selectivity for pitch chroma suggests that, at the level of cortex, the pitch map is circular rather than monotonic. For instance, it might resemble the pinwheel map of image orientation in visual cortex, where adjacent orientations are arranged like spokes around a central point (Bonhoeffer and Grinvald 1991). Circularity would ensure that the map is locally smooth, that is, nearby neurons share similar response properties. Local smoothness represents a key principle in the formation of cortical sensory maps (Swindale 1996).

Neurons that are selective for pitch chroma might underlie the perception of melody in music. Interconnection between different chroma-selective neurons might create sensitivity to common musical intervals. Evidence for such sensitivity has been found in neurophysiological recordings from the cat and monkey auditory cortex (Brosch et al. 1999; Brosch and Schreiner 2000).

The pitch adaptation paradigm developed in this study could be used to investigate the neural correlates of amusia. Congenital amusics make up about 4% of the general population. They have sometimes severe, lifelong difficulties in appreciating and producing music (Kalmus and Fry 1980). The causes of amusia are still a subject of debate (see Peretz and Hyde 2003, for review). The current paradigm yields a direct measure of neural pitch representation, unconfounded by task requirements, and might thus provide a tool for investigating whether amusia stems from a problem with the basic representation of pitch as opposed to the more general cognitive processes involved in music perception.

## Notes

*Conflict of Interest*: None declared.

## References

Abel C, Kössl M. 2009. Sensitive response to low-frequency cochlear distortion products in the auditory midbrain. J Neurophysiol. 101:1560–1574.

Barker D, Plack CJ, Hall DA. 2012. Reexamining the evidence for a pitch-selective region: a human fMRI study using iterated ripple noise. Cereb Cortex. 22:745–753.

Barth DS, Di S. 1990. Three-dimensional analysis of auditory-evoked potentials in rat neocortex. J Neurophysiol. 64:1527–1536.

Barth DS, Kithas J, Di S. 1993. Anatomic organization of evoked potentials in rat parietotemporal cortex: somatosensory and auditory responses. J Neurophysiol. 69:1837–1849.

Baumann S, Griffiths TD, Sun L, Petkov CI, Thiele A, Rees A. 2011. Orthogonal representation of sound dimensions in the primate midbrain. Nat Neurosci. 14:423–425.

Bell AJ, Sejnowski TJ. 1995. An information-maximization approach to blind separation and blind deconvolution. Neurol Comput. 7:1129–1159.

Bendor D, Wang X. 2010. Neural coding of periodicity in marmoset auditory cortex. J Neurophysiol. 103:1809–1822.

Bendor D, Wang X. 2005. The neuronal representation of pitch in primate auditory cortex. Nature. 436:1161–1165.

Bizley JF, Walker KMM, Silverman BW, King AJ, Schnupp JWH. 2009. Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. J Neurosci. 29:2064–2075.

Blakemore C, Campbell FW. 1969. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J Physiol. 203:237–260.

Bonhoeffer T, Grinvald A. 1991. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. Nature. 353:429–431.

Brosch M, Schreiner CE. 2000. Sequence sensitivity of neurons in cat primary auditory cortex. Cereb Cortex. 10:1155–1167.

Brosch M, Schulz A, Scheich H. 1999. Processing of sound sequences in macaque auditory cortex: response enhancement. J Neurophysiol. 82:1542–1559.

Butler RA. 1972. The auditory evoked response to stimuli producing periodicity pitch. Psychophysiol. 9:233–237.

Butler RA. 1968. Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. J Acoust Soc Am. 44:945–950.

Carlyon RP. 2004. How the brain separates sounds. Trends Cogn Sci. 8:465–471.

Cowan N, Winkler I, Teder W, Näätänen R. 1993. Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). J Exp Psychol. 19:909–921.

Crowley KE, Colrain IM. 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin Neurophysiol. 115:732–744.

Darwin CJ. 1997. Auditory grouping. Trends Cogn Sci. 9:327–333.

Delorme A, Makeig S. 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Meth. 134:9–21.

Dowling WJ. 1999. The development of music perception and cognition. In: Deutsch D, editor. The psychology of music. San Diego (CA): Academic Press. p. 603–625.

Efron B, Tibshirani RJ. 1993. An introduction to the bootstrap. New York, NY: Chapman & Hall.

Fishman YI, Reser DH, Arezzo JC, Steinschneider M. 1998. Pitch vs. spectral encoding of harmonic complex tones in primary auditory cortex of the awake monkey. Brain Res. 786:18–30.

Formisano E, Kim D-S, Di Salle F, van de Moortele P-F, Ugurbil K, Goebel R. 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron. 40:859–869.

Garrido MI, Kilner JM, Kiebel SJ, Friston KJ. 2007. Evoked brain responses are generated by feedback loops. Proc Natl Acad Sci U S A. 104:20961–20966.

Glasberg BR, Moore BCJ. 1990. Derivation of auditory filter shapes from notched-noise data. Hear Res. 47:103–138.

Grill-Spector K, Henson R, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Neurosci. 10:14–23.

Grill-Spector K, Malach R. 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. 107:293–321.

Gutschalk A, Patterson RD, Rupp A, Uppenkamp S, Scherg M. 2002. Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. NeuroImage. 15:207–216.

Gutschalk A, Patterson RD, Scherg M, Uppenkamp S, Rupp A. 2004. Temporal dynamics of pitch in human auditory cortex. NeuroImage. 22:755–766.

Hall DA, Edmondson-Jones AM, Fridriksson J. 2006. Periodicity and frequency coding in human auditory cortex. Eur J Neurosci. 24:3601–3610.

Hall DA, Johnsrude IS, Haggard MP, Palmer AR, Akeroyd MA, Summerfield AQ. 2002. Spectral and temporal processing in human auditory cortex. Cereb Cortex. 12:140–149.

Hubel DH, Wiesel TN. 1977. Functional architecture of macaque monkey visual cortex. Proc R Soc Lond. 198:1–59.

Jasper HH. 1958. The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol. 10:371–375.

Kalmus H, Fry DB. 1980. On tune deafness (dysmelodia): frequency, development, genetics and musical background. Ann Hum Genet. 43:369–382.

Kay RH, Matthews DR. 1972. On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulation tones. J Physiol. 225:657–677.

Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhöner B. 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb Cortex. 13:765–772.

Lee TW, Girolami M, Sejnowski TJ. 1999. Independent component analysis using an extended Infomax algorithm for mixed subgaussian and supergaussian sources. Neural Comput. 11:417–441.

Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ. 1997. Blind separation of auditory event-related brain responses into independent components. Proc Natl Acad Sci USA. 94:10979–10984.

McAlpine D. 2004. Neural sensitivity to periodicity in the inferior colliculus: evidence for the role of cochlear distortions. J Neurophysiol. 92:1295–1311.

McDermott JH, Lehr AJ, Oxenham AJ. 2010. Individual differences reveal the basis of consonance. Curr Biol. 20:1035–1041.

Megela AL, Teyler TJ. 1979. Habituation and the human evoked potential. J Comp Physiol Psychol. 6:1154–1170.

Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K. 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. NeuroImage. 13:684–701.

Näätänen R, Picton T. 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiol. 24:375–425.

Näätänen R, Sams M, Alho K, Paavilainen P, Reinikainen K, Sokolov EN. 1988. Frequency and location specificity of the human vertex N1 wave. Electroen Clin Neuro. 69:523–531.

Nelken I, Bizley JK, Nodal FR, Ahmed B, King AJ. 2008. Responses of auditory cortex to complex stimuli: functional organization revealed using intrinsic optical signals. J Neurophysiol. 99:1928–1941.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths T. 2002. The processing of temporal pitch and melody information in auditory cortex. Neuron. 36:767–776.

Penagos H, Melcher JR, Oxenham AJ. 2004. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J Neurosci. 24:6810–6815.

Peretz I, Hyde KL. 2003. What is specific to music processing? Insights from congenital amusia. Trends Cogn Sci. 7:362–367.

Puschmann S, Uppenkamp S, Kollmeier B, Thiel CM. 2010. Dichotic pitch activates pitch processing centre in Heschl's gyrus. NeuroImage. 49:1641–1649.

Rauschecker JP, Tian B, Hauser M. 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. Science. 268:111–114.

Sams M, Alho K, Näätänen R. 1983. Sequential effects on the ERP in discriminating two stimuli. Biol Psychol. 17:41–58.

Schellenberg EG, Trehub SE. 1994. Frequency ratios and the perception of tone patterns. Psychon Bull Rev. 1:191–201.

Scherer KR. 1995. Expression of emotion in voice and music. J Voice. 9:235–248.

Scherg M. 1990. Fundamentals of dipole source potential analysis. In: Grandori F, Hoke M, Romani G, editors. Auditory evoked potentials and fields. Basel, Switzerland: Karger. p. 40–69.

Schnupp JWH, Bizley JK. 2010. On pitch, the ear and the brain of the beholder. Focus on "Neural coding of periodicity in marmoset auditory cortex". J Neurophysiol. 103:1708–1711.

Schreiner CE, Cynader MS. 1984. Basic functional organization of second auditory cortical field (AII) of the cat. J Neurophysiol. 51:1284–1305.

Schwarz DWF, Tomlinson RWW. 1990. Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*). J Neurophysiol. 64:282–298.

Shackleton TM, Carlyon RP. 1994. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. J Acoust Soc Am. 95:3529–3540.

Smith DRR, Patterson RD. 2005. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. J Acoust Soc Am. 118:3177–3186.

Steinmann I, Gutschalk A. 2012. Sustained BOLD and theta activity in auditory cortex are related to slow stimulus fluctuations rather than to pitch. J Neurophysiol. 107:3458–3467.

Steinschneider M, Reser DH, Fishman YI, Schroeder CE, Arezzo JC. 1998. Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. J Acoust Soc Am. 104:2935–2955.

Swindale NV. 1996. The development of topography in the visual cortex: a review of models. Netw Comput Neural Syst. 7:161–247.

Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM. 2004. Tonotopic organization in human auditory cortex revealed by progressions of frequency selectivity. J Neurophysiol. 91:1282–1296.

Ueda K, Ohgushi K. 1987. Perceptual components of pitch: spatial representation using a multidimensional scaling technique. J Acoust Soc Am. 82:1193–1200.

Wacogne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S. 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci U S A. 108:20754–20759.

Warren JD, Uppenkamp S, Patterson RD, Griffiths TD. 2003. Separating pitch chroma and pitch height in the human brain. Proc Natl Acad Sci U S A. 100:10038–10042.

Wessinger CM, VanMeter J, Tian B, Lare JV, Pekar J, Rauschecker JP. 2001. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J Cogn Neurosci. 13:1–7.

Winkler I, Karmos G, Näätänen R. 1996. Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. Brain Res. 742:239–252.

Yost WA, Patterson R, Sheft S. 1996. A time domain description for the pitch strength of iterated rippled noise. J Acoust Soc Am. 99:1066–1078.