The effect of stimulus context on pitch representations in the human auditory cortex

Daphne Garcia [a,b,*], Deborah A. Hall [a,c], Christopher J. Plack [b]

[a] MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK
[b] Division of Human Communication and Deafness, The University of Manchester, Manchester M13 9PL, UK
[c] Division of Psychology, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK

ARTICLE INFO

ABSTRACT

Neuroimaging studies of pitch coding seek to identify pitch-related responses separate from responses to other properties of the stimulus, such as its energy onset, and other general aspects of the listening context. The current study reports the first attempt to evaluate these modulatory influences using functional magnetic resonance imaging (fMRI) measures of cortical pitch representations. Stimulus context was manipulated using a 'classical stimulation paradigm' (whereby successive pitch stimuli were separated by gaps of silence) and a 'continuous stimulation paradigm' (whereby successive pitch stimuli were interspersed with noise to maintain a stable envelope). Pitch responses were measured for two types of pitch-evoking stimuli; a harmonic-complex tone and a complex Huggins pitch. Results for a group of 15 normally hearing listeners revealed that context effects were mostly observed in primary auditory regions, while the most significant pitch responses were localized to posterior nonprimary auditory cortex, specifically planum temporale. Sensitivity to pitch was greater for the continuous stimulation conditions perhaps because they better controlled for concurrent responses to the noise energy onset and reduced the potential problem of a non-linear fMRI response becoming saturated. These results provide support for hierarchical processing within human auditory cortex, with some parts of primary auditory cortex engaged by general auditory energy, some parts of planum temporale specifically responsible for representing pitch information and adjacent regions that are responsible for complex higher-level auditory processing such as representing pitch information as a function of listening context.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Pitch is an important feature of auditory perception. It is arguably the most important perceptual feature of music and is a key component of tonal languages used in many parts of the world, such as sub-Saharan Africa and East Asia. In non-tonal languages such as English, we use pitch to recognize the gender and identity of different speakers as well as using intonation to discriminate between different types of sentence (e.g. a question or a statement, Chatterjee and Peng, 2008) and as a cue to stress. Pitch is also one of the main cues used by the auditory system to segregate sounds from different sources (Singh, 1987). Most pitch stimuli, whether natural or laboratory-made, are 'complex' tones made up of a number of harmonic sinusoidal components with frequencies that are integer multiples of the repetition rate or fundamental frequency (f0). In an early psychophysical study on the frequency analytical power of the human ear, Plomp (1964) discovered that the human ear is capable of 'hearing out' the first five to eight harmonics of a complex harmonic tone. These are the harmonics that each excite a different place on the basilar membrane, and are said to be 'resolved'. The basilar membrane can be modeled as a bank of bandpass filters, with a width corresponding to about 12% of the center frequency, for frequencies between 750 and 5000 Hz (Moore, 2003). Resolved harmonics fall within individual filters so that pitch may be determined by the distinctive pattern of spectral peaks in the neural excitation pattern. In contrast, for unresolved harmonics, multiple harmonics excite the same filter. For these stimuli, the pitch can be determined from the waveform produced by the interaction of the harmonics, whose repetition rate corresponds to the f0 of the complex tone (Houtsma and Smurzynski, 1990; Carlyon et al., 1992; Micheyl and Oxenham, 2004).

Although most pitch-evoking stimuli encountered in the environment are harmonic-complex tones, a pitch sensation can be evoked by manipulating noise signals, for example amplitude or frequency modulation (Mahaffey, 1967; Darwin et al., 1994) and spectral rippling (Yost and Hill, 1979). Cramer and Huggins (1958) found that pitch can even be conveyed through binaural interaction, with signals that contain no spectral or temporal pitch information when played individually to each ear. They presented the same wideband noise to both ears, except for a narrow frequency band, which was out of phase between the ears. A pitch was heard corresponding to the

* Corresponding author. MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK. Fax: +44 115 951 8503.
E-mail address: d.garcia@ihr.mrc.ac.uk (D. Garcia).

center frequency of the band. This 'Huggins pitch' (HP) is one of a number of binaural pitches that have now been identified (Plack and Oxenham, 2005).

Over recent years, neuroimaging methods such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG) and positron emission tomography (PET) have been used to search for the neural substrates of pitch processing in human listeners. These studies do not necessarily claim that pitch is first extracted in the auditory cortex, instead they simply seek to demonstrate that pitch is one of the organizing principles of sound coding at the level of the auditory cortex. Nevertheless, there are a number of discrepancies in the neuroimaging literature, especially in terms of neural mechanisms for pitch coding and the localization of those neural representations. Some authors have suggested that the same cortical neurons that represent pitch information are also involved in coding other aspects of sound, such as energy onset (Näätänen and Picton, 1987), while others claim that these properties are processed separately (Schönwiesner and Zatorre, 2008). Some authors have proposed that since a similar pitch percept can be elicited by sounds that possess very different spectral, temporal, and/or binaural characteristics, there should be a unified representation of pitch (Hall and Plack, 2009), while others argue that the physiological support for this claim is rather weak (Nelken et al., 2008).

A number of human neuroimaging studies have identified a pitch-sensitive region in the auditory cortex and have localized it to Heschl's gyrus (HG, see Fig. 3) (Griffiths et al., 1998; Gutschalk et al., 2002; Patterson et al., 2002; Hall et al., 2006; Puschmann et al., 2010). A number of these studies have favored the use of a single type of pitch-evoking stimulus known as iterated ripple noise (IRN). IRN is created by generating a sample of noise and imposing a delay before adding (or subtracting) the noise back to (or from) the original. The pitch sensation of the resulting sound is related to the reciprocal of the delay, and its salience is determined by the number of delay-and-add (or subtract) iterations and the gain applied to the delayed sample (Yost, 1996). One of the earliest studies to localize pitch representations was a PET study that identified areas of human auditory cortex that were sensitive to pitch salience (Griffiths et al., 1998). Here, salience was manipulated by systematically increasing the temporal regularity of IRN signals; with 0, 1, 2, 4, 8 and 16 iterations, respectively. Around HG, bilateral pitch-related activity was found to increase in magnitude with increasing pitch salience. Despite the data smoothing applied to the PET images, the focus of activity in the right hemisphere appeared to be close to the central portion of HG, while the focus in the left hemisphere appeared to be centered on lateral HG. Consistent with this finding, some MEG studies have implicated lateral HG in pitch processing by using click trains (Gutschalk et al., 2002, 2004, 2007). Furthermore, these findings are consistent with a non-human primate model of pitch coding that localizes pitch-selective neurons to a discrete cortical region near the anterolateral border of the primary auditory cortex (Bendor and Wang, 2005). This low-frequency region is proposed to correspond to lateral HG in humans.

Other neuroimaging studies have benefited from the greater spatial specificity of fMRI. For example, Hall et al. (2006) confirmed that pitch-related activity was present in lateral HG and tended to overlap with a primary-like region that was sensitive to low-frequency tones, irrespective of the spectral content of the (IRN) pitch-evoking stimuli. Results from a number of studies agree that the pitch-sensitive response is not confined to lateral HG, but spreads into adjacent posterior or anterior regions of the superior temporal gyrus (Patterson et al., 2002; Barrett and Hall, 2006; Penagos et al., 2004).

Hall and Plack (2009) have called into question the assumption that lateral HG operates as the main center for the cortical representation for pitch. Hall and Plack argued that evidence from one type of pitch-evoking stimulus alone does not constitute reliable evidence for a 'pitch center'. The motivation for their study was therefore to examine whether pitch-related responses in lateral HG

were consistently present for a range of different pitch-evoking stimuli, each with different physical characteristics. Pitch-evoking stimuli included IRN, single-frequency tones, wideband complex tones, missing f0 complex tones containing resolved or unresolved harmonics, and an HP stimulus. While IRN generated a pitch-sensitive response in lateral HG, the other pitch-evoking stimuli were more likely to produce activity in planum temporale (PT) than in lateral HG. The authors concluded that there was insufficient consistency across pitch effects to label any one region a 'pitch center'. However, these conclusions have recently been questioned by findings from an fMRI study that obtained significant responses in lateral HG and PT for two different HP stimuli and for a single-frequency tone-in-noise signal (Puschmann et al., 2010).

Typically, fMRI studies of pitch processing favor the presentation of a sequence of pitch-evoking sounds, each separated by silent intervals. This 'classical stimulation paradigm' is preferred because a slow repetition rate of stimulus bursts is known to evoke a robust and sustained fMRI response in auditory cortex (Harms and Melcher, 2002). Pitch-related activation is computed by subtracting from this condition the response to a baseline condition containing a matched sequence of noise bursts (Friston et al., 1996). Such subtraction methods rely on the assumption that the context of the stimulus presentation has no effect on the magnitude of the pitch-related response. In other subject areas, such as language processing (Price et al., 1997), this assumption has been shown to be untrue.

With regard to pitch, there is some evidence that auditory evoked responses are sensitive to the abrupt onset of sound energy (the energy-onset response) as well as to the abrupt onset of pitch (the pitch-onset response) (e.g. Krumbholz et al., 2003; Chait et al., 2006; Seither-Priesler et al., 2004). Thus, it is plausible that the pattern of energy onsets in the stimulus sequence might modulate the pitch-related response in a context-dependent manner.

The relationship between pitch- and energy-onset auditory evoked responses has been examined in detail in human listeners using MEG. The benefit of using MEG is that, unlike fMRI, it has millisecond temporal resolution that allows for the reliable detection of individual transient deflections (for a review see König et al., 2007). Krumbholz et al. (2003) separated the evoked response to the energy onset from that to the pitch onset by using a continuous stimulation paradigm. Here, the stimulus has a fixed spectral energy, but the perceptual features alternate between noise and pitch. The authors observed a transient deflection at about 150 ms after the transition from noise to pitch (the pitch-onset response), but not from pitch to noise, nor from one sample of noise to another. Moreover, the amplitude of the pitch-onset response increased as a function of pitch salience (number of iterations) and the latency of the pitch-onset response decreased as a function of pitch value (IRN delay). These results confirmed to the authors that the observed response was not simply related to detecting a perceptual change in the stimulus, but was indicative of pitch-specific coding. The study by Krumbholz et al. (2003) exclusively measured IRN, but similar properties of the pitch-onset response have been reported for a tone-in-noise stimulus and for HP (Chait et al., 2006).

Results from a recent depth-electrode study by Schönwiesner and Zatorre (2008) extend these findings from surface recordings of electromagnetic activity. The patient's pattern of brain activity revealed a double dissociation between the pitch-onset response and the energy-onset response. The former stimulated electrodes placed across lateral portions of HG while the latter stimulated electrodes placed across medial portions of HG. These results refute the idea that the same, or overlapping, populations of auditory cortical neurons respond to energy and pitch onsets.

Although the relatively poor temporal resolution of fMRI does not allow for the identification of individual evoked responses, we suggest that fMRI activation represents the accumulated activity resulting from a sequence of transient responses. Thus, the experiment described here constitutes the first attempt to investigate the differential

consequences of energy and pitch responses on fMRI measures of brain activity.

A difference between EEG and MEG measures of pitch coding and those of fMRI concerns the degree to which the response that is measured saturates at the upper limits of the response function. The fMRI response is known to be highly susceptible to non-linearities (Sidtis et al., 1999; Friston et al., 2000; Devor et al., 2003). Hence, the response to a sound stimulus that contains a combination of response-evoking features (e.g. energy and pitch onsets) will be most likely to exhibit saturation. If the fMRI response to a noise stimulus is brought close to saturation by the repeated onset of acoustical energy, any additional response (i.e. the addition of a pitch) will be limited by the saturation of the fMRI signal. This is illustrated in Fig. 1a. If the non-linear response model is correct, then the pitch-related activity (pitch condition minus noise condition) might be expected to be greater in the noise context than in the silent context because the former comparison is less affected by the saturating upper limit. This model has been invoked to explain previous auditory fMRI results (Melcher et al., 2000). The alternative model that proposes a linear system (e.g. Dale and Buckner, 1997) would predict an additive rather than a sub-additive response and this would be reflected in an equivalent pattern of pitch-related activation, irrespective of the stimulus context. The predictions of the linear model are illustrated in Fig. 1b.

In the present study, the energy-onset response was manipulated by presenting a sequence of pitch-evoking signals either within a silent context (akin to a 'classical stimulation' paradigm, see Hall and Plack, 2009) or a noise context (akin to the 'continuous stimulation' paradigm, see Krumbholz et al., 2003). For the silent context, we assume that the onset of each pitch will evoke both energy- and pitch-onset responses. For the noise context, we assume that the onset of each pitch will evoke only a pitch-onset response. We assume that sustained responses to pitch (see Gutschalk et al., 2004, 2007) are not markedly affected by the stimulus context and so do not contribute to any observed differences in pitch-related activity. To ensure the findings were not specific to a particular stimulus, the hypothesis was examined using two different pitch-evoking stimuli; an unresolved harmonic-complex tone (UNRES) and a complex HP (cHP).

Fig. 1. Two models depicting the coupling between neural activity and fMRI activation. (a) The non-linear response model suggests that the fMRI response is limited by a saturation level (dotted line) which, in the silent context condition, is dominated by multiple energy onsets so that the addition of a pitch elicits little additional activation. (b) The linear response model suggests that the response is additive. In this case fMRI activation is identical in silent and noise contexts.

Materials and methods

Listeners

Fifteen listeners (8 male, 7 female; age range 23–48 years) with normal hearing (≤20 dB hearing level between 250 Hz and 8 kHz) took part in this study. All but one listener (#01) was right-handed (laterality index = 50, Oldfield, 1971). Seven listeners were musically trained between grade 3 and diploma level (# 01, 02, 07, 08, 10, 12 and 15) while five others reported informal musical experience (self-taught/ungraded, # 04, 05, 09, 13 and 16). One listener (#11) completed the psychophysical testing but was not able to return for the fMRI session. None had a history of any neurological or hearing impairment. Listeners gave written informed consent and the study was approved by the Medical School Research Ethics Committee, University of Nottingham.

Stimuli

All stimuli evoked a pitch corresponding to a 200-Hz tone. One stimulus was a (diotic) unresolved harmonic-complex tone with harmonics 10–20 (henceforth referred to as UNRES). The level of each harmonic was 23 dB greater than the spectrum level of the control noise so that the gross spectral density of all the stimuli was the same. The UNRES stimulus was filtered between 2 and 4 kHz with a noise masker (49 dB SPL spectrum level) from 0 to 2 kHz (to mask cochlear distortion products). The other stimulus was a (dichotic) complex HP (henceforth referred to as cHP) in which the pitch cue was only available via integration of the signals from each ear (dichotic). The cHP stimulus was created from a diotic Gaussian noise (49 dB SPL spectrum level) filtered between 0 and 4 kHz. In one ear, a π phase shift was introduced in eight 30-Hz wide frequency bands, centered on the first eight harmonics. The noise control stimulus was a Gaussian noise (49 dB SPL spectrum level), again low-pass filtered at 4 kHz. The three signals (UNRES, cHP and noise) were matched in bandwidth (0–4 kHz) and spectral density (and hence overall energy). It is probably impossible to generate stimuli that differ in pitch strength but are perfectly matched for every other perceptual feature. While the Gaussian noise is a good control for cHP (the only other perceptual difference between the two is the spatiality of the decorrelated band), it is perhaps less so for UNRES because the signals differ in other respects, such as their envelope structure. The logic of 'common activity' has been applied to reduce the risk of attributing these potential differences to pitch (Hall and Plack, 2009; Puschmann et al., 2010) and can also be applied to the analysis of the current experiment.

For the psychophysical testing, each pitch stimulus was 350 ms in duration (including 10-ms linear-intensity onset and offset ramps) and the inter-stimulus interval was 500 ms. Reference stimuli had an f0 of 200 Hz. The stimuli were presented at an overall level of 85 dB SPL, calibrated using a KEMAR manikin (Burkhard and Sachs, 1975) fitted with Bruel and Kjaer half-inch microphone type 4134 (serial no. 906663), Zwislocki occluded ear simulator (Knowles model no. DB-100) and Bruel and Kjaer measuring amplifier type 2636 (serial no. 1324093), scaled from 22.4 Hz to 22.4 kHz using fast time constant (125 ms) on maximum hold. Due to the metallic components in the KEMAR system, calibration inside the scanner was not possible.

In the scanner, stimulus conditions each comprised a 15.41-s alternating sequence of 450-ms experimental sounds each separated by 230 ms. In the 'pitch-in-noise-context' conditions, the separation contained a Gaussian noise as the context. In the 'pitch-in-silent-context' conditions, the pitch signals were separated by 230 ms silence. The first and last components of each sequence were the context. Each pitch and noise signal was generated using 10 ms linear-intensity onset and offset ramps, which were overlapped at the 3 dB points to produce a stable envelope for the stimulus (see Fig. 2).
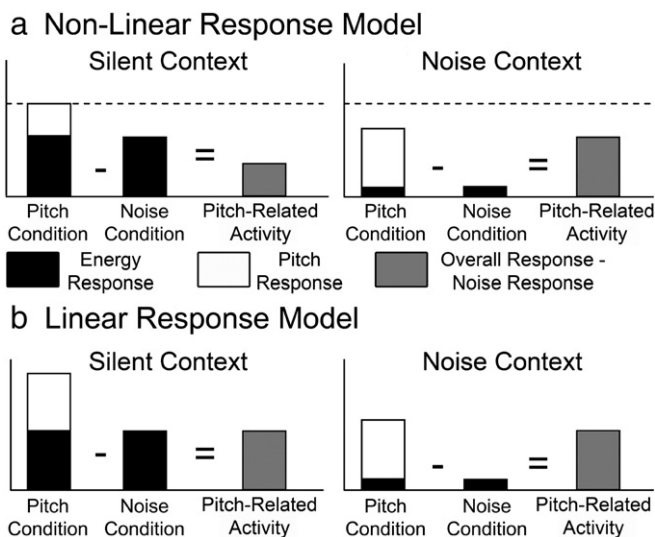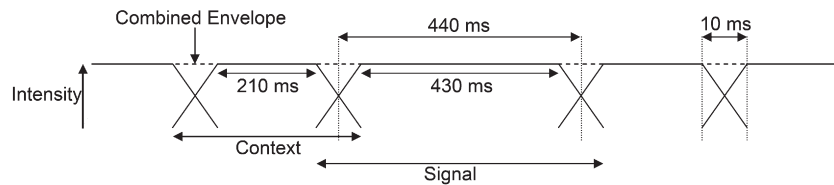
Fig. 2. Schematic diagram of the signal and context components of the stimuli, overlapped to produce a stable envelope. In the 'pitch-in-noise-context' and 'pitch-in-silent-context' conditions, the signal is either UNRES or cHP and the context is Gaussian noise or silence, respectively. In the 'noise-in-noise-context' condition, both signal and context segments are Gaussian noise, hence it is a continuous noise. In the 'noise-in-silent-context' condition, the signal is Gaussian noise and the context is silence.

Eighteen sample sequences were created for each condition. The control noise conditions were created in the same way.

Psychophysical testing

Prior to the scanning session, each participant performed a pitch-discrimination test to assess accuracy in distinguishing the pitch cues. Psychophysical testing was carried out in a sound-attenuating booth and stimuli were delivered through Sennheiser HD 480 II headphones. Stimuli were presented through custom-made software that is supported by the Matlab platform (The MathWorks, Natick, MA). Pitch-discrimination thresholds were measured for cHP and UNRES using a three alternative forced-choice, two-down, one-up, adaptive procedure that targeted 70.7% performance (Levitt, 1971). Two observation intervals contained the standard tone (UNRES or cHP) with an f0 of 200 Hz. The remaining interval (chosen at random) contained a comparison tone with a higher f0 which the listener was required to select as the 'odd one out'. On the first trial, the f0 difference between standard and comparison was 20% (40 Hz). The percent difference increased or decreased by a factor of two for the first four reversals, and by a factor of 1.414 for the final 12 reversals. Discrimination threshold was taken as the geometric mean of the f0 difference at the final 12 reversals. The adaptive track was not allowed to increase above 200% (600 Hz). Responses were recorded and stored electronically. On each trial, feedback was given via a green (correct) or red (incorrect) light on the software interface. There were five runs each for cHP and UNRES; the first was considered as practice and so the pitch-discrimination threshold was taken as the average of the last four runs.

fMRI protocol

Scanning was performed on a Philips 3 T Intera Acheiva using an 8-channel SENSE receiver head coil. A T1-weighted high-resolution (1 mm^3) anatomical image (matrix size = 256 × 256, 160 saggital slices, TR = 8.2 ms, TE = 3.7 ms) was collected for each subject. The anatomical scan was used to position the functional scan centrally on HG, and care was taken to include the entire superior temporal gyrus and to exclude the eyes. Functional scanning used a T2*-weighted echo-planar sequence with a voxel size of 3 mm^3 (matrix size = 64 × 64, 32 oblique-axial slices, TE = 36 ms). Sparse imaging with a TR of 8000 ms and a clustered acquisition time of 1969 ms was used (Edmister et al., 1999; Hall et al., 1999). A SENSE factor of 2 was applied to reduce image distortions and a SofTone factor of 2 was used to reduce the background scanner noise level by 9 dB. Functional data were acquired over two runs of 128 scans each, with the sounds presented in a quasi-random order, and with the rule that the same stimulus condition was not presented twice in succession. Listeners were requested to listen to the sounds, but were not required to perform any task. A custom-built MR compatible system delivered distortion-free sound using high-quality electrostatic headphones (Sennheiser HE60 with high-voltage amplifier HEV70) that had been specifically modified for use during fMRI.

Data analysis

Images were analyzed separately for each listener using statistical parametric mapping (SPM5, http://www.fil.ion.ucl.ac.uk/spm). Pre-processing steps included realignment to correct for subject motion, normalization of individual scans to a standard image template, and smoothing with a Gaussian filter of 8 mm full width at half maximum. Individual analyses were computed for the two runs (256 scans), specifying the two pitch and the two noise conditions as separate regressors in the design. In the individual analysis, we specified separate statistical contrasts for each sound condition relative to the silent baseline that was implicitly modeled in the design. A high-pass filter cutoff of 420 s was used.

First, the data for individual participants was analyzed using a first-level general linear model to assess the effects of interest with respect to the scan-to-scan variability. The resulting model estimated the fit of the design matrix (X) to the data (Y) in each voxel in order to provide β values (the contribution of a single regressor to the overall fMRI signal). In order to obtain activation maps for individual analysis, SPM was used to fit the GLM to each individual voxel in the functional image, and to compute individual t statistics. The effect of each stimulus condition was identified and the resulting (unthresholded) contrast images were entered into a group-level random effects analysis in order to assess the effects of interest with respect to the inter-subject variability. At this group level, 2 × 2 repeated measures ANOVAs were created, with signal (pitch present and pitch absent, i.e. noise) and context (noise and silent contexts) as factors. Separate ANOVAs were computed for UNRES and cHP conditions and within each ANOVA, simple main effects and interactions were calculated using t statistics (Friston et al., 2005). Although initial SPM t contrasts were defined using an uncorrected threshold of $p < 0.001$, all results are reported after small volume correction (SVC) to control for type I errors using a false discovery rate (FDR) threshold of $p < 0.05$ (Genovese et al., 2002). The small volume defined the auditory cortex across the superior temporal gyrus (including HG, PT and planum polare) and contained 4719 voxels in the left hemisphere and 5983 voxels in the right hemisphere. Activations were localized using an SPM toolbox that overlays an SPM thresholded map onto a set of probabilistic maps of the three cyto-architectonic subdivisions of HG (Te 1.0, Te 1.1 and Te 1.2, Morosan et al., 2001; Eickhoff et al., 2006).

Results

Behavioral results

The mean geometric discrimination threshold across the listeners for cHP was 2.93 Hz and for UNRES was 3.54 Hz. A paired t-test showed that thresholds for the two stimuli did not differ significantly [$t(1,14) = 0.053$, $p = 0.821$]. For comparison, previous research suggests that the threshold for a 200-Hz pure tone is ~1 Hz (Wier et al., 1977), and that for a 200-Hz unresolved harmonic-complex tone is ~5 Hz (Houtsma and Smurzynski, 1990).
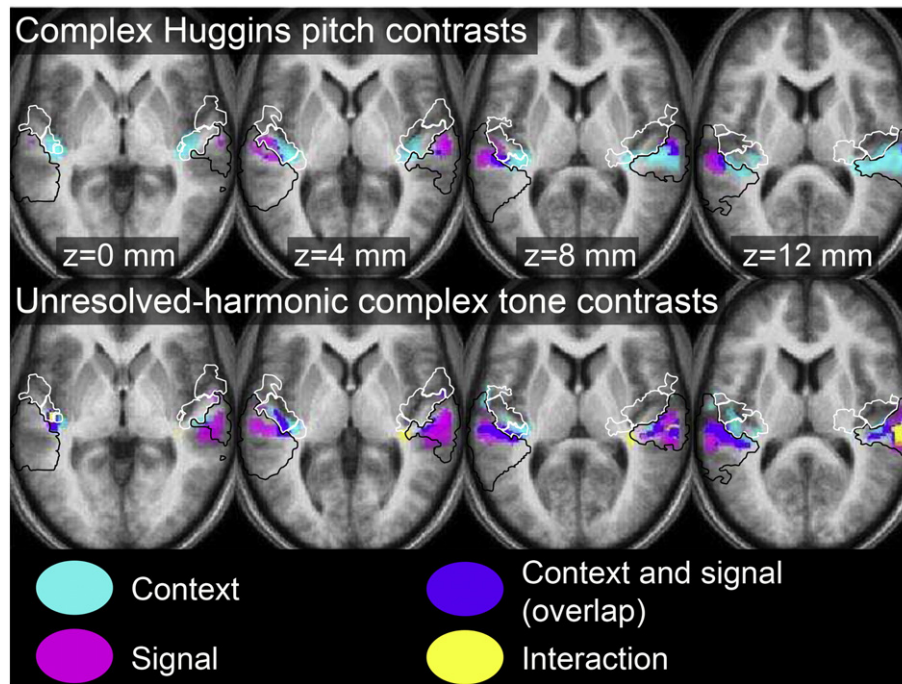
Fig. 3. Activation map from the 2×factorial ANOVA showing locations for the main effects of context (cyan) and signal (magenta), regions where the two main effects overlap (purple), and areas in which context modulates pitch (yellow). The white borders denote areas Te 1.1 (medial portion), Te 1.0 (middle portion) and Te 1.2 (lateral portion) (Morosan et al., 2001) on Heschl's gyrus. The black border outlines PT (Westbury et al., 1999). Activation is overlaid onto an average anatomical image made from the 15 individual listeners. The left hemisphere is on the left-hand side of each anatomical image.

fMRI results: effect of stimulus context

For the main effect of context, bilateral clusters of activation ($p<0.05$ SVC) were revealed for both pitch types (UNRES and cHP, Fig. 3). These are regions in which activation was greater for the silent context than for the noise context. Both UNRES and cHP showed the same pattern of context-related activation. In the left hemisphere, the most significant activation was in the medial portion of HG (Te 1.1, Morosan et al., 2001) (x −38 y −26 z 6 mm for both pitch types,

Table 1). In the right hemisphere, the most significant activation was in PT (cHP: x 64 y −30 z 12 mm; UNRES: x 66 y −22 z 10 mm). However, for both pitch types there was a substantial spread of context-related activation across bilateral HG (areas Te 1.0, central HG; Te 1.1, medial HG, and Te 1.2, lateral HG) and PT (Fig. 3). There were no voxels that showed a greater response for the noise context than the silent context, which suggests that the human auditory cortex is more responsive to successive energy onsets than it is to the overall energy in the stimulus.

Table 1
Significant clusters of activity for cHP and UNRES contrasts. The peak voxels of activity are reported for the left and right hemispheres, respectively.

	Left hemisphere					Right hemisphere				
	Peak coordinates	Z-score	Voxel-level p-value[a]	Cluster size	Location	Peak coordinates	Z-score	Voxel-level p-value[a]	Cluster size	Location
cHP contrasts										
Context: silence>noise	−38 −26 6	5.79	<0.001	1181	Te 1.1	64 −30 12	5.36	<0.001	1182	PT
	−66 −42 20	3.39	0	8	PT	–	–	–	–	–
	−42 −14 20	3.34	0	23	Te 1.0	–	–	–	–	–
	−44 −20 14	3.18	0.01	2	Te 1.0	–	–	–	–	–
Pitch: cHP>noise	−58 −24 8	4.48	0	314	PT	64 −16 6	5.5	<0.001	137	PT
	−48 −6 −8	3.22	0.02	3	PP	56 −36 4	3.09	0.02	1	PT
Interaction	No suprathreshold voxels					No suprathreshold voxels				
UNRES contrasts										
Context: silence>noise	−38 −26 6	5.28	<0.001	1375	Te 1.1	66 −22 10	4.49	0	550	PT
	–	–	–	–	–	62 −4 8	4.01	0	29	Te 1.0
	−60 −38 14	3.43	0	24	PT	66 −36 24	3.35	0.01	5	PT
	−36 −18 16	3.26	0.01	2	Te 1.1	38 −20 −2	3.18	0.01		Te 1.1
Pitch: UNRES>noise	−62 −24 8	4.91	<0.001	856	PT	66 −18 6	5.7	<0.001	746	PT
	−46 −8 −6	3.35	0	29	Te 1.0	62 −4 2	3.86	0	22	Te 1.2
	−60 −12 4	3.35	0	14	Te 1.2	46 −14 −8	3.21	0.01	4	PP
Interaction	−64 −40 20	3.56	0.04	20	PT	36 −30 4	4.19	0.04	45	PT
	−46 −18 −2	3.47	0.04	16	Te 1.0	68 −30 12	4.11	0.04	234	PT
	−44 −38 20	3.19	0.05	1	PT	42 −40 16	3.24	0.05	2	PT

[a] FDR-corrected.

fMRI results: effect of pitch

The effect of signal (pitch present versus pitch absent) was also computed from the 2×2 full factorial ANOVA to identify auditory cortical regions in which activation was greater for the pitch condition than for the spectrally matched noise control. Both cHP and UNRES contrasts revealed large bilateral clusters of pitch-related activation (Table 1). For both pitch types, the greatest response was located bilaterally in PT. Peak voxels were within 4 mm of each other ($x -58$ $y -24$ z 8 mm and $x -62$ $y -24$ z 8 mm in the left hemisphere for cHP and UNRES respectively; and again x 64 $y -16$ z 6 mm and x 66 $y -18$ z 6 mm in the right). While it is true that the pitch-related activity for UNRES appeared somewhat more widespread than that for cHP, there was an extremely high agreement between the most significant peaks (Table 1). This finding increases our confidence that this activation focus represents a response to the pitch quality of the stimuli, and not to some other feature that was not perfectly matched between conditions. There were no voxels that showed a significantly greater response to the control noise than to the pitch stimulus.

According to a number of slices displayed in Fig. 3, some of the pitch-related activity appears to span Heschl's sulcus (the posterior border of HG) and so we explored the data further to establish where the central focus of activity was located in relation to this landmark, as well as in relation to the lateral HG response reported by Puschmann et al. (2010), in Table 2 of their paper. The spatial coordinate of the most significant response to cHP was extracted for each individual listener and these data were used to compute a mean coordinate and its 95% confidence intervals in each dimension. The cHP contrast was chosen as it was most comparable to the HP stimulus reported by Puschmann et al. Panel a (Fig. 4) shows that the peaks were separated by 11 and 8 mm in the left and right hemispheres, respectively, with our focus being posterior to that of Puschmann et al. (2010) on the posterior side of Heschl's sulcus. It is perhaps also worth noting here that our result is not entirely contradictory with the neuroimaging literature. Indeed, even Puschmann et al. (2010) observed some significant bilateral pitch-related activity in PT. The anterolateral portion of PT has been widely associated with the representation of nonspatial auditory features, as shown by a meta-analysis (Arnott et al., 2004). This is illustrated in Fig. 4b.

Regional differences in the response to context and pitch

The activation maps generated from the SPM analysis indicated a preference for stimulus context in HG and a preference for pitch in PT and lack of any clear preference for either feature in lateral HG. To quantify these putative differences between regions, we conducted a number of region-of-interest analyses enabling direct statistical comparison between regions. Three spatially discrete regions were defined using the probabilistic values for areas Te 1.0 and Te 1.2 (Morosan et al., 2001), and for PT (Westbury et al., 1999). Using all voxels within each region-of-interest, the mean estimate of the size of the fMRI response to each stimulus was computed separately for each listener. For UNRES and cHP, repeated measures ANOVAs were specified to examine differential responses to the stimulus context across Te 1.0 and PT. For both types of pitch, there was a significant interaction between context and region [$F(1,26) = 17.53, p<0.001$ for UNRES and $F(1,26) = 18.43, p<0.001$ for cHP]. While Te 1.0 was highly sensitive to stimulus context (Fig. 5a), PT was significantly less so (Fig. 5c).

Fig. 5c also demonstrates how PT is sensitive to UNRES and cHP. For both contrasts, the main effect of pitch within PT reached significance [$F(1,26) = 38.76, p<0.001$ for UNRES and $F(1,26) = 10.03$, $p = 0.004$ for cHP]. In Te 1.2, there was support for a preference for UNRES compared to the control noise [$F(1,26) = 9.72, p = 0.004$], but this was not true for cHP ($p = 0.146$) (see Fig. 5b). To examine whether these apparent regional differences were significant, we
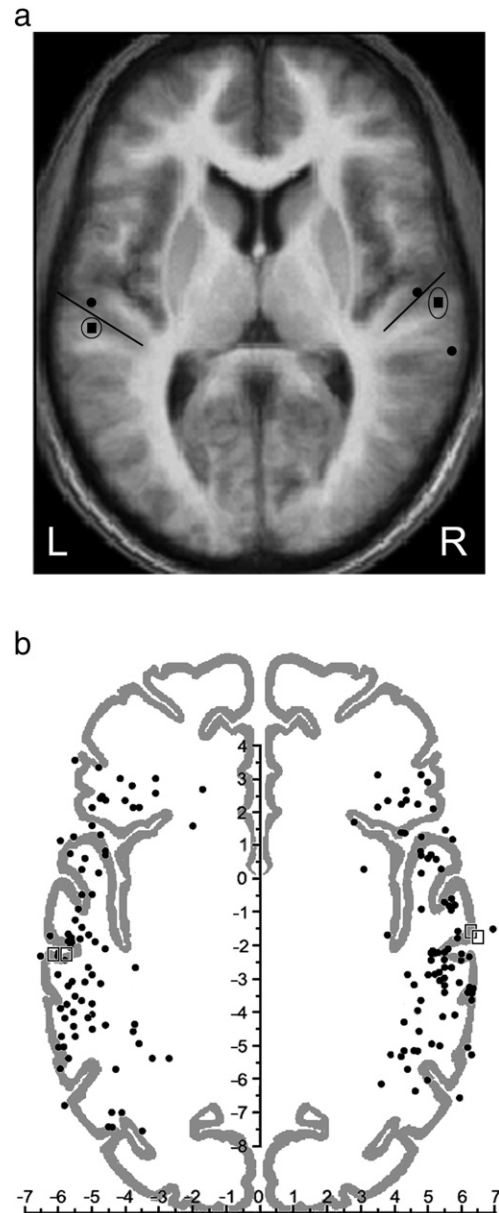

Fig. 4. (a) The peak locations of pitch-related activity in the current study (black squares) and their counterparts reported by Puschmann et al. (2010) (black circles). The two oblique black lines represent the posterior border of Heschl's gyrus (Heschl's sulcus). The underlying brain image is the mean normalized anatomical scan for our group of 15 listeners. (b) A schematic axial view ($z = -4$ mm) denoting the coordinates of nonspatial auditory activity (black dots) plotted on the corresponding outline of the Talairach brain (using data reported by Arnott et al., 2004). Equivalent data for the pitch contrasts reported in Table 1 are overlaid onto the same image (black squares), after a linear transformation to convert the coordinate space appropriately (using the procedure reported by Arnott et al., 2004).

performed a direct comparison between Te 1.2 and PT, again using ANOVA statistics. The overall effect of pitch was significant [$F(1,26) = 18.82, p<0.001$ for UNRES and $F(1,26) = 4.58, p = 0.042$ for cHP] but there were no significant interactions between pitch and region for either stimulus type. In other words, the effect of pitch was not significantly greater in PT than in Te 1.2. It should be noted that the estimates of pitch-related activity in PT are rather conservative because the estimates of response magnitude were averaged over a large number of voxels, and many of those in posterior PT were unresponsive to any of the stimulus features. Thus, taken together, these analyses indicate a clear pitch-related response in PT, but they do not rule out the possibility of a similar pitch response in Te 1.2 (i.e., lateral HG).
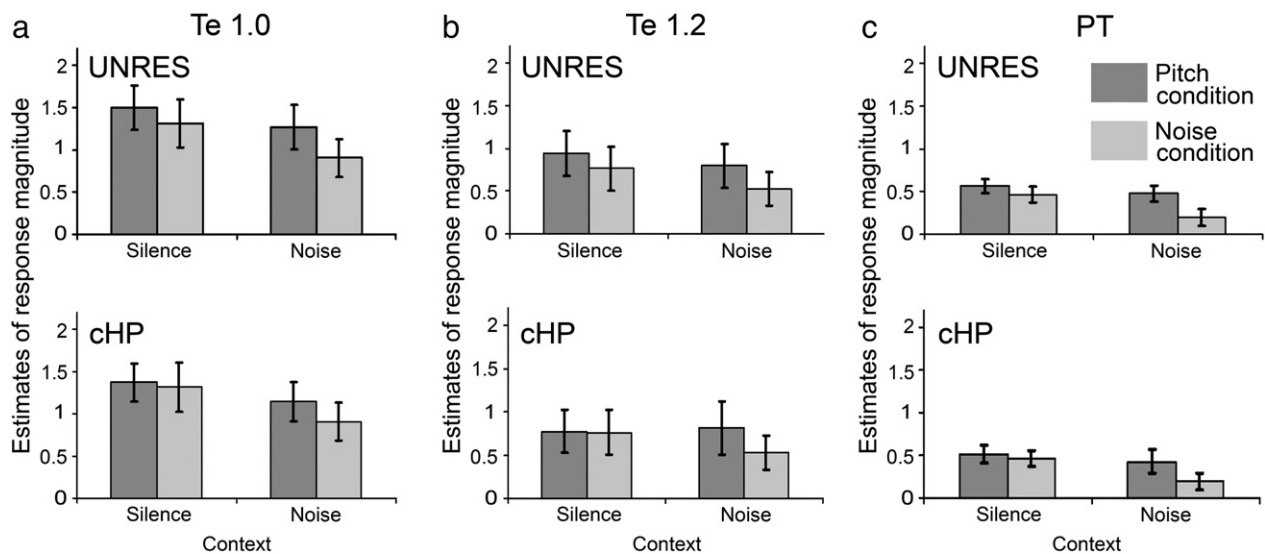
Fig. 5. Graphical representation of the effects of context and pitch, and their interactions plotted separately for UNRES and cHP. The three panels represent the three different regions of interest: (a) primary auditory cortex defined by Te 1.0, (b) lateral HG defined by Te 1.2, and (c) PT. Error bars represent 95% confidence intervals around the estimated mean activation.

Interaction showing modulatory effect of context on pitch-related activity

Fig. 3 illustrates some overlap between the effects of context and pitch. The interaction term from the factorial ANOVA was examined to determine the pattern of co-activation because a significant interaction would demonstrate that the pitch-related activity was modulated by the stimulus context. Cortical regions showing a significant interaction between pitch and context are shown in yellow in Fig. 3 and are reported in Table 1. For the UNRES conditions, a number of small foci of bilateral activity were located in PT and one in left central HG (Te 1.0). No significant interaction was observed for the cHP conditions, although there was some evidence for a similar distribution of activity at the uncorrected threshold ($p < 0.001$, not shown).

To understand the shape of the interaction, a post-hoc region-of-interest analysis was computed again using the estimates of size of the response in PT. As Fig. 5c shows, although the responses in the noise context were marginally smaller than in the silent context, the difference between the pitch and noise conditions was much more marked in the noise context. Post-hoc testing demonstrated that this difference reached significance ($p < 0.05$) in the UNRES condition.

Three observations are consistent with the non-linear model; (i) the region is sensitive to detecting changes in energy and in pitch, but the two effects are not additive, (ii) the larger responses in the silent context than in the noise context are slight but at least consistent with the interpretation that the accumulated response to successive energy onsets contributed to the overall magnitude of the observed activity, and (iii) the significantly greater pitch-related activity in the noise context than in the silent context would be expected if the former comparison was less affected by the saturating upper limit.

Discussion

The present fMRI study reports a novel attempt to measure the effects of stimulus context on the cortical representation of pitch. Concurring with previous findings from the same authors (Hall and Plack, 2009), the most significant pitch-related activity was centered on posterior auditory cortex, in lateral PT. The results are consistent with the view that posterior auditory cortex is engaged in nonspatial, as well as in spatial, auditory analysis (Arnott et al., 2004). Some pitch-related activity was identified in lateral HG (Te 1.2). However, there was no convincing evidence for a general sensitivity to pitch in

this region because the effect was only significant for UNRES. The UNRES contrast is not ideally controlled for non-pitch features such as temporal modulations in the UNRES signal that are not present in the control noise.

Focal subdivisions of PT revealed a modulatory effect such that the magnitude of the pitch response was determined by some higher-order property of the stimulus, which was determined by a particular combination of features (i.e. pitch and context). A reasonable conjecture is that the pattern of results within these subdivisions of PT is consistent with the (non-linear response) model; in which the same neural population is responsive to both energy onsets and pitch onsets and in which the sum of those responses has a maximum saturating limit. Non-linearity is perhaps not restricted to the fMRI methodology since neuromagnetic studies of the pitch-onset response have also reported a greater sensitivity to pitch when the energy-onset response has been eliminated by presenting the pitch signals in a noise context rather than in a silent context (Krumbholz et al., 2003; Seither-Priesler et al., 2004; Chait et al., 2006). These MEG studies have convincingly demonstrated that these pitch-onset effects cannot be attributed to a general response to stimulus change. Furthermore, if the context effects seen here were simply a *non-specific* response to stimulus alternation, one would expect that the subtraction of the continuous noise condition from the pitch-in-noise conditions would elicit the equivalent pattern of activity as the subtraction of the continuous noise from the noise-in-silence condition. This was not the case.

Although the response to energy onsets was greatest within primary auditory cortex (including Te 1.0, Fig. 5a), this effect was by no means restricted to primary regions. Given the widespread effects of stimulus context, it is cautionary to note that previous fMRI reports of pitch-related activity could have perhaps underestimated the cortical representations of pitch whenever those experiments utilized the 'classical stimulation' paradigm. For example, we speculate that this explanation might account for the rather low consistency of pitch-related activity reported for the 16 listeners in Hall and Plack (2009).

fMRI evidence for a hierarchy of auditory processing across primary and nonprimary regions

The results from the present fMRI study indicate a hierarchy of auditory processing where physical properties related to the temporal

structure of the acoustic energy engage HG (primary auditory cortex) and perceptual features such as pitch dominate the response in PT (nonprimary auditory cortex). In addition, subdivisions of PT appear sensitive to particular combinations of features (i.e. context and pitch). It is possible to speculate that the observed effects of context and pitch represent a modulation of stimulus processing by the changing stimulus features involuntarily capturing the focus of selective attention, and do not isolate stimulus-driven processing. Two reasons lead us to believe this not to be the case. First, a carefully designed fMRI study seeking to measure the effects of pitch-related attention in the auditory cortex by manipulating the focus of attention towards or away from the pitch of an IRN stimulus, failed to find any such effects (Krumbholz et al., 2007, see also Altmann et al., 2008; Paltoglou et al., 2009). Second, it is unclear why attentional capture by changes in pitch or changes in stimulus energy might engage different parts of the auditory cortex, as observed in the current set of results.

With respect to the present localization of the energy-onset response in human primary auditory cortex, it is perhaps worthwhile returning to the findings from the depth-electrode study reported by Schönwiesner and Zatorre (2008). Recall, they reported energy-onset responses recorded at an electrode placed on medial HG and pitch-onset responses at an electrode placed on lateral HG, suggesting spatially separate neural populations responsive to the two sound attributes. Their results are somewhat consistent with the hierarchical model that we propose to account for our fMRI data. Certainly, we would agree that medial HG is strongly responsive to energy onsets because this was the main activation site where the fMRI response was significantly diminished when energy onsets were eliminated by the noise context. In the present study, there was partial evidence for the engagement of lateral HG in the response to pitch. However, the present fMRI data would predict that if one is to place an electrode array intracranially across the surface of PT instead of across HG, then an even greater pitch response might be recorded.

It is generally accepted that 'higher' cortical regions encode perceptual qualities, such as spatial location or speech sound identity (Davis and Johnsrude, 2003; Nelken, 2008; Rauschecker et al., 1995; Rauschecker, 1998). The human nonprimary region PT would fit into this conceptualization. However, we do not claim that the role of primary auditory cortex is merely to encode simple (i.e. physical) attributes of a stimulus. In fact, it would perhaps be rather naïve to do so. For example, electrophysiological recordings have established that primary auditory cortical neurons have complex response properties, showing sensitivity to both low-level and high-level features of sounds (Nelken, 2008; Kelly and Sally, 1988). Nevertheless, the present study demonstrates a reasonably clear division between HG and PT in terms of representations of context and pitch, respectively. Moreover, only subdivisions of PT (the 'higher' cortical region) were sensitive to response interactions between stimulus context and pitch representations.

Pitch-related activity in planum temporale (PT), for both UNRES and cHP

The location of pitch-related activity found in the present factorial fMRI experiment is more posterior than previously suggested by fMRI studies that have used IRN as the pitch-evoking stimulus (e.g. Patterson et al., 2002; Hall et al., 2005), have applied dipole source modeling to MEG data in order to localize the neural generators of the pitch response (Krumbholz et al., 2003; Gutschalk et al., 2002, 2004, 2007) or have applied a region-of-interest mask defining lateral HG (Puschmann et al., 2010). Our data demonstrate that the *most reliable* location of pitch-related activity appears to be immediately behind lateral HG, in PT.

To ensure the findings were not specific to a particular stimulus, the effect of context on pitch representations was examined using two different pitch-evoking stimuli. The finding that the location of responses to UNRES and cHP are broadly comparable suggests that the pitch responsive region observed in the present study is not stimulus specific. However, UNRES elicited activation in a greater number of voxels. Although inter-listener consistency was low in the Hall and Plack (2009) study, an incidence map of their 16 individual listeners also demonstrates similar patterns for their unresolved harmonic complex and simple HP. Information on the number of activated voxels is not available for that study, but they did not find a significant difference in percentage signal change for the two conditions. In the present study, most of the pitch-related activity was centered on PT in accordance with Hall and Plack's (2009) findings. Future investigations using a greater number of different pitch-evoking stimuli would be required to determine whether the results reported here are general to all pitch stimuli or are specific to certain types of pitch stimuli.

References

Altmann, C.F., Henning, M., Döring, M.K., Kaiser, J., 2008. Effects of feature-selective attention on auditory pattern and location processing. NeuroImage 41, 69–79.

Arnott, S.R., Binns, M.A., Grady, C.L., Alain, C., 2004. Assessing the auditory dual-pathway model in humans. NeuroImage 22, 401–408.

Barrett, D.J.K., Hall, D.A., 2006. Response preferences for "what" and "where" in human non-primary auditory cortex. NeuroImage 32, 968–977.

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

Burkhard, M.D., Sachs, R.M., 1975. Anthropometric manikin for acoustic research. J. Acoust. Soc. Am. 58, 214–222.

Carlyon, R.P., Demany, L., Semal, C., 1992. Detection of across-frequency differences in fundamental frequency. J. Acoust. Soc. Am. 91, 279–292.

Chait, M., Poeppel, D., Simon, J.Z., 2006. Neural response correlates of detection of monaurally and binaurally created pitches in humans. Cereb. Cortex 16, 835–848.

Chatterjee, M., Peng, S.-C., 2008. Processing F0 with cochlear implants: modulation frequency discrimination and speech intonation recognition. Hear. Res. 235, 143–156.

Cramer, E.M., Huggins, W.H., 1958. Creation of pitch through binaural interaction. J. Acoust. Soc. Am. 30, 413–417.

Dale, A.M., Buckner, R.L., 1997. Selective averaging of rapidly presented individual trials using fMRI. Hum. Brain Mapp. 5, 329–340.

Darwin, C.J., Ciocca, V., Sandell, G.J., 1994. Effects of frequency and amplitude-modulation on the pitch of a complex tone with a mistuned harmonic. J. Acoust. Soc. Am. 95, 2631–2636.

Davis, M.H., Johnsrude, I.S., 2003. Hierarchical processing in spoken language comprehension. J. Neurosci. 23, 3423–3431.

Devor, A., Dunn, A.K., Andermann, M.L., Ulbert, I., Boas, D.A., Dale, A.M., 2003. Coupling of total hemoglobin concentration, oxygenation, and neural activity in rat somatosensory cortex. Neuron 39, 353–359.

Edmister, W.B., Talavage, T.M., Ledden, P.J., Weiskoff, R.M., 1999. Improved auditory cortex imaging using clustered volume acquisitions. Hum. Brain Mapp. 7, 89–97.

Eickhoff, S.B., Amunts, K., Mohlberg, H., Zilles, K., 2006. The human parietal operculum. II. Steretaxic maps and correlation with functional imaging results. Cerebral Cortex 16, 268–279.

Friston, K.J., Price, C.J., Fletcher, P., Moore, C., Frackowiak, R.S.J., Dolan, R.J., 1996. The trouble with cognitive subtraction. NeuroImage 4, 97–104.

Friston, K.J., Mechelli, A., Turner, R., Price, C.J., 2000. Nonlinear responses in fMRI: the balloon model, Volterra kernels, and other hemodynamics. NeuroImage 12, 466–477.

Friston, K.J., Stephan, K.E., Lund, T.E., Morcom, A., Kiebel, S., 2005. Mixed-effects and fMRI studies. NeuroImage 24, 244–252.

Genovese, C., Lazar, N., Nichols, T., 2002. Thresholding of statistical maps in neuroimaging using the false discovery rate. NeuroImage 15, 870–878.

Griffiths, T.D., Büchel, C., Frackowiak, R.S.J., Patterson, R.D., 1998. Analysis of temporal structure in sound by the human brain. Nat. Neurosci. 1, 422–427.

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

Gutschalk, A., Patterson, R.D., Scherg, M., Uppenkamp, S., Rupp, A., 2007. The effect of temporal context on the sustained pitch response in human auditory cortex. Cereb. Cortex 17, 552–561.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., et al., 1999. "Sparse" temporal sampling in auditory fMRI. Hum. Brain Mapp. 7, 213–223.

Hall, D.A., Barrett, D.J.K., Akeroyd, M.A., Summerfield, A.Q., 2005. Cortical representations of temporal structure in sound. J. Neurophysiol. 94, 3181–3191.

Hall, D.A., Edmondson-Jones, A.M., Fridriksson, J., 2006. Frequency and periodicity coding in human auditory cortex. Eur. J. NeuroSci. 24, 3601–3610.

Hall, D.A., Plack, C.J., 2009. Pitch processing sites in the human auditory brain. Cereb. Cortex 19, 576–585.

Harms, M.P., Melcher, J.R., 2002. Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. J. Neurophysiol. 88, 1433–1450.

Houtsma, A.J.M., Smurzynski, J., 1990. Pitch identification and discrimination for complex tones with many harmonics. J. Acoust. Soc. Am. 87, 304–310.

Kelly, J.B., Sally, S.L., 1988. Organization of auditory cortex in the albino rat: binaural response properties. J. Neurophysiol. 59, 1756–1769.

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

Krumbholz, K., Eickhoff, S.B., Fink, G.R., 2007. Feature- and object-based attentional modulation in the human auditory 'where' pathway. J. Cogn. Neurosci. 19, 1721–1733.

König, R., Sieluzycki, C., Durka, P.J., 2007. Tiny signals from the human brain: acquisition and processing of biomagnetic fields in magnetoencephalography. J. Low Temp. Phys. 146, 697–718.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.

Mahaffey, R.B., 1967. Some effects of cosine amplitude modulation on pitch of tone bursts. J. Acoust. Soc. Am. 41, 1593.

Melcher, J.R., Sigalovsky, I.S., Guinan Jr., J.J., Levine, R.A., 2000. Lateralized tinnitus studied with functional magnetic resonance imaging: abnormal inferior colliculus activation. J. Neurophysiol. 83, 1058–1072.

Micheyl, C., Oxenham, A.J., 2004. Sequential F0 comparisons between resolved and unresolved harmonics: no evidence for translation noise between two pitch mechanisms. J. Acoust. Soc. Am. 116, 3038–3050.

Moore, B.C.J., 2003. An Introduction to the Psychology of Hearing, 5th ed. Elsevier, London.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial references 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. Psychophysiology 24, 375–425.

Nelken, I., Bizley, J.K., Nodal, F.R., Ahmed, B., King, A.J., Schnupp, J.W.H., 2008. Responses of auditory cortex to complex stimuli: functional organization revealed using intrinsic optical signals. J. Neurophysiol. 99, 1928–1941.

Nelken, I., 2008. Processing of complex sounds in the auditory system. Curr. Opin. Neurobiol. 18, 413–417.

Oldfield, R.C., 1971. Assessment and analysis of handedness — Edinburgh inventory. Neuropsychologia 9, 97–113.

Paltoglou, E., Sumner, C.J., Hall, D.A., 2009. Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention. Hear. Res. 257, 106–118.

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

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

Plack, C.J., Oxenham, A.J., 2005. The psychophysics of pitch. In: Plack, C.J., Oxenham, A.J., Fay, R.R., Popper, A.N. (Eds.), Pitch: Neural Coding and Perception. Springer, New York.

Plomp, R., 1964. The ear as a frequency analyzer. J. Acoust. Soc. Am. 36, 1628–1636.

Price, C.J., Moore, C.J., Friston, K.J., 1997. Subtractions, conjunctions, and interactions in experimental design of activation studies. Hum. Brain Mapp. 5, 264–272.

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

Rauschecker, J.P., 1998. Cortical processing of complex sounds. Curr. Opin. Neurobiol. 8, 516–521.

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

Schönwiesner, M., Zatorre, R.J., 2008. Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. Exp. Brain Res. 187, 97–105.

Seither-Priesler, A., Krumbholz, K., Patterson, R., Seither, S., Lütkenhöner, B., 2004. Interaction between the neuromagnetic responses to sound energy onset and pitch onset suggests common generators. Eur. J. NeuroSci. 19, 3073–3080.

Sidtis, J.J., Strother, S.C., Anderson, J.R., Rottenberg, D.A., 1999. Are brain functions really additive? Neuroimage 9, 490–496.

Singh, P.G., 1987. Perceptual organization of complex-tone sequences: a tradeoff between pitch and timbre? J. Acoust. Soc. Am. 82, 886–899.

Westbury, C.F., Zatorre, R.J., Evans, A.C., 1999. Quantifying variability in the planum temporale: a probability map. Cereb. Cortex 9, 392–405.

Wier, C., Jesteadt, W., Green, D., 1977. Frequency discrimination as a function of frequency and sensation level. J. Acoust. Soc. Am. 61, 178–184.

Yost, W.A., Hill, R., 1979. Models of the pitch and pitch strength of ripple noise. J. Acoust. Soc. Am. 66, 400–410.

Yost, W., 1996. Pitch of iterated rippled noise. J. Acoust. Soc. Am. 100, 511–518.