

Elsevier Editorial System(tm) for NeuroImage
Manuscript Draft

Manuscript Number: NIMG-09-2259R1

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

Article Type: Regular Article

Section/Category: Systems Neuroscience

Corresponding Author: Miss Daphne Garcia,

Corresponding Author's Institution: MRC Institute of Hearing Research

First Author: Daphne Garcia

Order of Authors: Daphne Garcia; Deborah A Hall; Christopher J Plack

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 nonlinear 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.

The submitted manuscript considers pitch coding in the human brain. Pitch is arguably the most important feature of auditory perception because of its role in language and music perception. Over the years, a number of human neuroimaging studies have sought 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. However, most of these particular studies have used MEG to isolate the transient pitch-related onset response, rather than consider the ongoing response to the stimulus.

In this paper, we report a novel fMRI investigation of the effects of stimulus context on pitch representations in human auditory cortex. Our main finding was that of a hierarchy of auditory processing whereby areas of primary auditory cortex were mainly responsive to acoustical energy, pitch information was processed more posteriorly in areas of planum temporale and complex higher-level properties of sound (in this case, a context-dependent pitch response) were processed in adjacent areas of planum temporale. The finding that the pitch response is modulated by the stimulus context is important for future studies because it demonstrates how sensitivity to detecting pitch-related activity might be enhanced by presenting pitch stimuli from a noise context rather than from a silent context. We explain this result perhaps because it better controls for concurrent responses to the noise energy onset and reduces the potential problem of a nonlinear fMRI response becoming saturated.

Response to Reviewers :

Reviewer 1 considered the study to be of great interest and well designed. We were invited to expand on the apparent discrepancies between other studies specifically with respect to the lack of specific activity in lateral Heschl's gyrus. Reviewer 1 was helpful in making a number of suggestions for potentially resolving these differences by conducting some additional, exploratory analysis.

(1) One suggestion was for a region-of-interest in each of the three separate subdivisions of Heschl's gyrus (for comparison to PT) that could be presented as an addition to Figure 4.

We thank Reviewer 1 for this insightful comment. We have conducted region of interest analyses (for Te 1.0, Te 1.2 and PT) to specifically test for regional differences in the response to context and pitch. These results are reported in the revised manuscript and we have expanded Figure 5 (originally Figure 4) as recommended. The pattern of responses was equivalent in the two primary auditory regions (Te 1.0 and Te 1.1) and so Te 1.1 has not been reported because it offered no additional insight.

Pages 18-19 : ‘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 of 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 condition 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 (Figure 5a), PT was significantly less so (Figure 5c).

Figure 5c also demonstrates how PT is sensitive to both UNRES and cHP. For both contrasts, the main effect of pitch within PT reached significance [$F(1,26)=38.76$, $p<0.001$ and $F(1,26)=10.03$, $p=0.004$ for UNRES and cHP respectively]. In Te 1.2, there was some 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 Figure 5b). To examine whether these apparent regional differences were significant, we 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, so that the effect of pitch was not significantly greater in PT than in Te 1.2. 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. However, it should be noted that these 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.’

**(2) Two important papers should be discussed in some detail:
Bendor and Wang (2005) Nature 436:1161-1165
Puschmann et al (2010) Neuroimage 49:1641-1649**

We thank Reviewer 1 for bringing the 2010 paper by Puschmann to our attention and we have discussed both papers where appropriate.

Page 5 : ‘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 Figure 3) (Griffiths et al., 1998; Gutschalk et al., 2002; Patterson et al., 2002; Hall et al., 2006; Puschmann et al., 2010).’

Page 6 : ‘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.’

Page 7 : ‘These conclusions have recently been questioned by findings from an fMRI study that obtained significant responses in lateral HG and PT for two different binaural pitch-evoking stimuli and for a single frequency tone-in-noise signal (Puschmann et al., 2010).’

(3a) From Figure 3, the reviewer inferred that the pitch-specific activity appears pretty close to lateral HG. Hence, there may be only a little discrepancy with previous data (e.g. Puschmann et al 2010) since the discrepancy could simply be reduced to whether the area is at the anterior or posterior site of Heschl's gyrus.

We have addressed this by making a direct visual comparison between our data and that of Puschmann et al. using confidence intervals to demonstrate statistical reliability. A discussion of this comparison has been added to the Results section, and a figure has been included to show the two sets of peaks with respect to the posterior border of Heschl’s gyrus.

Pages 17-18 : ‘According to a number of slices displayed in Figure 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 co-ordinate 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 (Figure 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 Figure 4b.’

(3b) The specificity of the pairwise contrasts depends greatly on the matching of the pitch and the control stimulus. For the Huggins pitch, noise appears to be a good match. Complex tones and noise are different in several respects other than pitch saliency, including random fluctuations of the noise, different envelope structure, etc. The influence of such parameters needs to be discussed.

In our previous paper (Hall and Plack, 2009) we readily acknowledged the difficulty in perfectly matching stimuli for every feature except for their pitch. However, we also noted that the confounding variable is not always the same for each stimulus pair (modulations for complex tones and IRN, resolvable spectral features for resolved harmonics, slow spectrotemporal fluctuations for IRN, spatial location for Huggins pitch etc). In the 2009 paper, we argued that by locating the activity common to *all* stimulus contrasts, it is possible to maximize the chance of finding the site that responds selectively to pitch. The same logic of interpretation was applied by Puschmann et al. (2010) and can also be applied here. While it is true that the pitch-related activity for the complex-tone contrast was somewhat more widespread than that for the Huggins-pitch contrast, there was an extremely high agreement at the peaks of the cluster. Table 1 demonstrates that the main peaks for the two pitch types differ by only 4 mm in either hemisphere. This finding increases our confidence that this activation focus represents a response to the pitch quality of the stimuli and not to some other parameter.

A version of the above discussion has been added on pages 11 and 17.

(4) The context effect and the interaction in Figure 4 might not necessarily be related to pitch specificity and hence require explanation by the saturation model. The signals compared here are one where two stimuli are alternated with one where a continuous stimulus runs through all time (the alternation with silence will produce a similar sequence pattern). My guess was that a similar pattern was found if the complex tone was continuous instead of the noise, such that the larger difference may simply reflect the presence versus absence of stimulus alternation.

Whichever way the problem is posed, the main point that we wish to make is that the response in this region is essentially non-linear. So for example, if a brain region was sensitive to detecting changes in energy and in pitch, then the two effects are not additive. This point has been clarified in the text.

We had noted in the Introduction on page 8 that MEG studies have shown that the pitch-onset response occurs 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. This is strong support for it being a pitch-specific response.

We discuss this issue further in the Discussion on page 22.

Page 22 : ‘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.'

(5) Logothetis 2001 is the wrong paper in the context of BOLD non-linearities. I'd rather cite it for a linear model. Devor et al 2003 might be a better match here.

We thank the Reviewer for this more appropriate reference and we have substituted Devor et al. (2003) on page 9.

(6) It was unclear to Reviewer 1 what we meant by the phrase "sustained (steady-state) responses" because they refer to two distinct entities in MEG research.

This was an error and so we have removed the term 'steady-state' from the text on page 10.

(7) In the discussion, the authors refer to Schönwiesener and Zatorre to suggest that had they used electrodes in Planum temporal, they might have found larger pitch responses there. This is clearly not what the MEG studies suggest, which sample activity from both, Planum temporale and Heschl's gyrus.

The paper by Schönwiesener and Zatorre reported the result of an invasive procedure in which pitch responses were recorded using an intracranial electrode array. By our comment, we were specifically making a prediction about a possible future set of intracranial recordings that could be made from planum temporale. Intracranial recordings have greater spatial precision than MEG and could make a valuable contribution to clarifying the localization of pitch selective responses in human auditory cortex. We have clarified this point by adding 'intracranially' to the text on page 21.

(8) Incorrect spelling of "Rauschecker" on page 27

This has been corrected.

Reviewer 2 raised a number of comments.

(9) The description of stimuli and conditions on page 11 was unclear. "The level of each harmonic was 23 dB greater than the spectrum level of noise". To which noise does this refer? What was the effective intensity of these noise stimuli?

The level of each harmonic was 23 dB greater than the spectrum level of the control noise. We have also clarified the details of the spectrum level of the pitch and noise signals which was 49 dB SPL in all cases.

(10) Were the control noise stimuli also presented with gaps and with noise context (filling the gaps) like the pitch stimuli?

Page 12 : ‘The control noise conditions were created in the same way.’

(11) What was the order of different conditions in the fMRI experiment?

Page 14 : ‘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.’

(12) The description of psychophysical testing needs to be elaborated. Were the reference tones with F0 of 200 Hz sinusoidal or harmonic tones? Were the reference tones and test tones presented in pairs, and what was their presentation order? What were the "three alternatives" in the forced-choice pitch discrimination task?

Page 13 : “Two observation intervals contained the standard tone (UNRES or CHP) with a 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).”

(13) With respect to the behavioral results, it is said that the mean discrimination threshold for CHP and UNRES tones was 2.93 and 3.54 Hz, respectively. What would be the threshold for sinusoidal 200-Hz tones or for harmonic tones with F0 of 200 Hz?

Page 16 : ‘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).’

Wier, C., Jesteadt, W., Green, D., 1977. Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America* 61, 178-184.

(14) In the present fMRI study, "listeners were requested to listen to the sounds, but were not required to perform any task" (p. 13). A problem with such 'passive listening' is that there is no way to know what the listeners were doing in the scanner. Were they attending to the sounds or ignoring them, or were they attending more to some sound than others? Possible effects of attention is a highly relevant issue because a number of studies, not cited by the authors, have shown effects of attention to pitch on auditory-cortex responses measured with fMRI (e.g., Hall et al *Hum Brain Mapp* 2000, Alain et al *PNAS* 2001, Petkov et al *Nat Neurosci* 2004, Degerman et al *Brain Res* 2006, Johnson & Zatorre *Neuroimage* 2006, Rinne et al *J Neurosci* 2009).

Reviewer 2 therefore wondered whether the pitch and context effects reported here might actually be effects of covert or involuntary attention to the eliciting sounds rather than genuine effects of pitch or energy change?

This is certainly an interesting consideration since attending to a visual stimulus feature, such as colour or motion, is well known to enhance the processing of that feature in the visual cortex (Corbetta et al., 1990, *Science*). There are a growing number of equivalent studies that examine the neural basis of the facilitatory role of attention in listening tasks that direct attention to stimulus features, such as pitch. Nevertheless, after carefully weighing up this evidence (see Paltoglou et al., 2009), we believe that the support for any selective enhancement of pitch-related fMRI activity is not very convincing. Furthermore, it is unclear why attentional capture by changes in pitch or changes in stimulus energy might engage different parts of the auditory cortex.

Page 20 : ‘It perhaps remains 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 iterate-ripple noise 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 the different parts of the auditory cortex observed in the current set of results.’

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.

Krumbholz, K., Eickhoff, S.B., Fink, G.R., 2007. Feature- and object-based attentional modulation in the human auditory ‘where’ pathway. *Journal of Cognitive Neuroscience* 19, 1721–1733.

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. *Hearing Research*, 257, 106–118.

Cited from p107, Paltoglou et al. : “Using fMRI to quantify auditory cortical activity in nine listeners, Petkov et al. (2004, *Nature Neurosci*) concluded that while responses in primary auditory cortex were tonotopically arranged, attentional enhancement occurred primarily in non-primary auditory cortex and was not frequency specific. This null result contradicts with other reports of significant feature-specific enhancement. For example, one fMRI study used a task that manipulated the focus of attention by instructing its nine listeners to discriminate either phonemes or spatial locations (Ahveninen et al., 2006, *PNAS*). The results showed enhancement in non-primary auditory regions posterior to Heschl’s gyrus (HG) when attending to the spatial feature, although there was no significant differential effect of attending to the nonspatial feature. For a group of 19 listeners, Krumbholz et al. (2007, *J Cog*

Neurosci) again showed a significant increase in activity when attention was directed to spatial motion compared to pitch in posterior nonprimary motion-sensitive areas, especially in the right temporoparietal junction. A recent fMRI study has also confirmed the same feature-specific ‘asymmetry’ for spatial and nonspatial selective attention (Altmann et al., 2008, NeuroImage). For a group of 12 listeners, the magnitude of feature-specific adaptation in motion-sensitive auditory cortical regions was influenced by whether attention was directed to location or to the spectrotemporal pattern, whereas adaptation in pattern-sensitive regions was not significantly modulated by the listening task.”

(15) Reviewer 2 invited us to consider the implications of our pitch-related activity in PT for the dual-pathways model in which PT is generally held to be responsible for spatial processing and anterior auditory regions responsible for non-spatial (e.g., pitch) processing (c.f. Arnott et al Neuroimage, 2004).

A pervasive model of functional specialization in the auditory cortex suggests two pathways that predominantly code sound identity and sound location information along anterior and posterior pathways, respectively. While this model provides a useful framework, it is sometimes misinterpreted that the function of the posterior nonprimary auditory regions is *exclusively* that of spatial analysis. Instead, the data support the view that planum temporale is engaged in a range of spatial and nonspatial tasks. For example, following their meta-analysis, Arnott et al. (2004) conclude that nonspatial activity was distributed throughout the temporal lobe. Temporal regions posterior to the primary auditory cortex were active in spatial studies (6 out of 11) and nonspatial studies (15 of 27). This is entirely consistent with our data presented in the current study.

Page 17-18 : ‘It is perhaps also worth observing that the most significant peaks of pitch-related activity fell within the lateral part of PT, as shown by Arnott et al. (2004) for the representation of other nonspatial auditory features (Figure 4).’

Page 21 : ‘Concurring with previous findings from the same authors (Hall and Plack, 2009), pitch-related activity was mostly centered on posterior auditory cortex, in lateral PT and is consistent with the view that posterior auditory cortex is engaged in nonspatial, as well as in spatial, auditory analysis (Arnott et al., 2004).’

We have included a new figure (Figure 4) which reports the distribution of activation peaks for nonspatial auditory tasks as defined by the meta-analysis published by Arnott et al. (2004). We have overlaid our peaks onto the same figure for direct comparison.

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

(16) Reviewer 2 asked us to reconsider our interpretation of the relative locations of activity across the different pitch signals and for the pitch and context conditions as reported on pages 16 and 17.

We thank Reviewer 2 for highlighting that the peak voxels for the effect of pitch signal were within 3-4 mm of each other, not 8 mm as was originally reported. This has now been corrected in the text.

Reviewer 2 recommended that we evaluate whether the peaks for the pitch response were at a significant distance from the context-related activation peaks? We agree that this is an important point since only a significant difference between peak locations would justify the suggestion that the pitch and context activated different areas of the auditory cortex.

This has been addressed in response to point (1) above.

(17) Reviewer 2 recommended a change to the last sentence in the abstract since the areas where the responses are generated may be sensitive to pitch information and listening context.

This has been rephrased to read ‘...adjacent regions that are responsible for complex higher-level auditory processing such as representing pitch information as a function of listening context.’

(18) "data was" should be "data were".

This has been corrected on page 13.

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

3

4 Daphne Garcia^{1,2*}

5 Deborah A. Hall^{1,3}

6 and

7 Christopher J. Plack²

8

9 ¹MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD,
10 UK

11

12 ²Division of Human Communication and Deafness, The University of
13 Manchester, Manchester M13 9PL, UK

14

15 ³Division of Psychology, Nottingham Trent University, Burton Street,
16 Nottingham NG1 4BU, UK

17

18 *Corresponding author

19

20 Telephone: +44 (0) 115 922 3431

21 Fax: +44 (0) 115 951 8503

22 Email: d.garcia@ihr.mrc.ac.uk

23

24

1 **Abstract**

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

25

1 **Introduction**

2 Pitch is an important feature of auditory perception. It is arguably the most
3 important perceptual feature of music and is a key component of tonal
4 languages used in many parts of the world, such as sub-Saharan Africa and
5 East Asia. In non-tonal languages such as English, we use pitch to recognize
6 the gender and identity of different speakers as well as using intonation to
7 discriminate between different types of sentence (e.g. a question or a
8 statement, Chatterjee and Peng, 2008) and as a cue to stress. Pitch is also
9 one of the main cues used by the auditory system to segregate sounds from
10 different sources (Singh, 1987). Most pitch stimuli, whether natural or
11 laboratory-made, are 'complex' tones made up of a number of harmonic
12 sinusoidal components with frequencies that are integer multiples of the
13 repetition rate or fundamental frequency (f_0). In an early psychophysical study
14 on the frequency analytical power of the human ear, Plomp (1964) discovered
15 that the human ear is capable of 'hearing out' the first five to eight harmonics
16 of a complex harmonic tone. These are the harmonics that each excite a
17 different place on the basilar membrane, and are said to be 'resolved'. The
18 basilar membrane can be modelled as a bank of bandpass filters, with a width
19 corresponding to about 12 per cent of the center frequency, for frequencies
20 between 750 and 5000 Hz (Moore, 2004). Resolved harmonics fall within
21 individual filters so that pitch may be determined by the distinctive pattern of
22 spectral peaks in the neural excitation pattern. In contrast, for unresolved
23 harmonics, multiple harmonics excite the same filter. For these stimuli, the
24 pitch can be determined from the waveform produced by the interaction of the
25 harmonics, whose repetition rate corresponds to the f_0 of the complex tone

1 (Houtsma and Smurzynski, 1990; Carlyon et al., 1992; Micheyl and Oxenham,
2 2004).

3 Although most pitch-evoking stimuli encountered in the environment
4 are harmonic complex tones, a pitch sensation can be evoked by
5 manipulating noise signals, for example amplitude or frequency modulation
6 (Mahaffey, 1967; Darwin, et al., 1994) and spectral rippling (Yost and Hill,
7 1979). Cramer and Huggins (1958) found that pitch can even be conveyed
8 through binaural interaction, with signals that contain no spectral or temporal
9 pitch information when played individually to each ear. They presented the
10 same wideband noise to both ears, except for a narrow frequency band, which
11 was out of phase between the ears. A pitch was heard corresponding to the
12 center frequency of the band. This 'Huggins pitch' (HP) is one of a number of
13 binaural pitches that have now been identified (Plack and Oxenham, 2005).

14 Over recent years, neuroimaging methods such as functional magnetic
15 resonance imaging (fMRI), electroencephalography (EEG),
16 magnetoencephalography (MEG) and positron emission tomography (PET)
17 have been used to search for the neural substrates of pitch processing in
18 human listeners. These studies do not necessarily claim that pitch is first
19 extracted in the auditory cortex, instead they simply seek to demonstrate that
20 pitch is one of the organizing principles of sound coding at the level of the
21 auditory cortex. Nevertheless, there are a number of discrepancies in the
22 neuroimaging literature, especially in terms of neural mechanisms for pitch
23 coding and the localization of those neural representations. Some authors
24 have suggested that the same cortical neurons that represent pitch
25 information are also involved in coding other aspects of sound, such as

1 energy onset (Näätänen and Picton, 1987), while others claim that these
2 properties are processed separately (Schönwiesner and Zatorre, 2008). Some
3 authors have proposed that since a similar pitch percept can be elicited by
4 sounds that possess very different spectral, temporal, and/or binaural
5 characteristics, there should be a unified representation of pitch (Hall and
6 Plack, 2009), while others argue that the physiological support for this claim is
7 rather weak (Nelken et al., 2008).

8 A number of human neuroimaging studies have identified a pitch
9 sensitive region in the auditory cortex and have localized it to Heschl's gyrus
10 (HG, see Figure 3) (Griffiths et al., 1998; Gutschalk et al., 2002; Patterson et
11 al., 2002; Hall et al., 2006; Puschmann et al., 2010). A number of these
12 studies have favored the use of a single type of pitch-evoking stimulus known
13 as iterated ripple noise (IRN). IRN is created by generating a sample of noise
14 and imposing a delay before adding (or subtracting) the noise back to (or
15 from) the original. The pitch sensation of the resulting sound is related to the
16 reciprocal of the delay, and its salience is determined by the number of delay-
17 and-add (or subtract) iterations and the gain applied to the delayed sample
18 (Yost, 1996). One of the earliest studies to localize pitch representations was
19 a PET study that identified areas of human auditory cortex that were sensitive
20 to pitch salience (Griffiths et al., 1998). Here, salience was manipulated by
21 systematically increasing the temporal regularity of IRN signals; with 0, 1, 2, 4,
22 8 and 16 iterations, respectively. Around HG, bilateral pitch-related activity
23 was found to increase in magnitude with increasing pitch salience. Despite the
24 data smoothing applied to the PET images, the focus of activity in the right
25 hemisphere appeared to be close to the central portion of HG, while the focus

1 in the left hemisphere appeared to be centered on lateral HG. Consistent with
2 this finding, some MEG studies have implicated lateral HG in pitch processing
3 by using click trains (Gutschalk et al., 2002; 2004; 2007). **Furthermore, these**
4 **findings are consistent with a non-human primate model of pitch coding that**
5 **localizes pitch-selective neurons to a discrete cortical region near the**
6 **anterolateral border of the primary auditory cortex (Bendor and Wang, 2005).**
7 **This low-frequency region is proposed to correspond to lateral HG in humans.**

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

16 Hall and Plack (2009) have called into question the assumption that
17 lateral HG operates as the main center for the cortical representation for pitch.
18 Hall and Plack argued that evidence from one type of pitch-evoking stimulus
19 alone does not constitute reliable evidence for a 'pitch center'. The motivation
20 for their study was therefore to examine whether pitch-related responses in
21 lateral HG were consistently present for a range of different pitch-evoking
22 stimuli, each with different physical characteristics. Pitch-evoking stimuli
23 included IRN, single frequency tones, wideband complex tones, missing f0
24 complex tones containing resolved or unresolved harmonics, and an HP
25 stimulus. While IRN generated a pitch-sensitive response in lateral HG, the

1 other pitch-evoking stimuli were more likely to produce activity in planum
2 temporale (PT) than in lateral HG. The authors concluded that there was
3 insufficient consistency across pitch effects to label any one region a 'pitch
4 center'. However, these conclusions have recently been questioned by
5 findings from an fMRI study that obtained significant responses in lateral HG
6 and PT for two different HP stimuli and for a single-frequency tone-in-noise
7 signal (Puschmann et al., 2010).

8 Typically, fMRI studies of pitch processing favor the presentation of a
9 sequence of pitch-evoking sounds, each separated by silent intervals. This
10 'classical stimulation paradigm' is preferred because a slow repetition rate of
11 stimulus bursts is known to evoke a robust and sustained fMRI response in
12 auditory cortex (Harms and Melcher, 2002). Pitch-related activation is
13 computed by subtracting from this condition the response to a baseline
14 condition containing a matched sequence of noise bursts (Friston et al.,
15 1996). Such subtraction methods rely on the assumption that the pitch cue is
16 the only difference between the two conditions. In other words, the context of
17 the stimulus presentation has no effect on the magnitude of the pitch-related
18 effect. In other subject areas, such as language processing (Price et al.,
19 1997), this assumption has been shown to be untrue.

20 With regard to pitch, there is some evidence that auditory evoked
21 responses are sensitive to the abrupt onset of sound energy (the energy-
22 onset response) as well as to the abrupt onset of pitch (the pitch-onset
23 response) (e.g. Krumbholz et al., 2003; Chait et al., 2006; Seither-Priesler et
24 al., 2004). Thus, it is plausible that the pattern of energy onsets in the stimulus

1 sequence might modulate the pitch-related response in a context-dependent
2 manner.

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

22 Results from a recent depth-electrode study by Schönwiesner and
23 Zatorre (2008) extend these findings from surface recordings of
24 electromagnetic activity. The patient's pattern of brain activity revealed a
25 double dissociation between the pitch-onset response and the energy-onset

1 response. The former stimulated electrodes placed over lateral portions of HG
2 while the latter stimulated electrodes placed over medial portions of HG.
3 These results refute the idea that the same, or overlapping, populations of
4 auditory cortical neurons respond to energy and pitch onsets.

5 Although the relatively poor temporal resolution of fMRI does not allow
6 for the identification of individual evoked responses, we suggest that fMRI
7 activation represents the accumulated activity resulting from a sequence of
8 transient responses. Thus, the experiment described here constitutes the first
9 attempt to investigate the differential consequences of energy and pitch
10 responses on fMRI measures of brain activity.

11 A difference between EEG and MEG measures of pitch coding and
12 those of fMRI concern the degree to which the response that is measured
13 saturates at the upper limits of the response function. The fMRI response is
14 known to be highly susceptible to non-linearities (Sidtis et al., 1999; Friston et
15 al., 2000; [Devor et al., 2003](#)). Hence, the response to a sound stimulus that
16 contains a combination of response-evoking features (e.g. energy and pitch
17 onsets) will be most likely to exhibit saturation. If the fMRI response to a noise
18 stimulus is brought close to saturation by the repeated onset of acoustical
19 energy, any additional response (i.e. the addition of a pitch) will be limited by
20 the saturation of the fMRI signal. This is illustrated in Figure 1A. If the non-
21 linear response model is correct, then the pitch-related activity (pitch condition
22 minus noise condition) might be expected to be greater in the noise context
23 than in the silent context because the former comparison is less affected by
24 the saturating upper limit. This model has been invoked to explain previous
25 auditory fMRI results (Melcher et al., 2000). The alternative model that

1 proposes a linear system (e.g. Dale and Buckner, 1997) would predict an
2 additive rather than a sub-additive response and this would be reflected in an
3 equivalent pattern of pitch-related activation, irrespective of the stimulus
4 context. The predictions of the linear model are illustrated in Figure 1B.

5 ** INSERT FIGURE 1 AROUND HERE **

6 In the present study, the energy-onset response was manipulated by
7 presenting a sequence of pitch-evoking signals either within a silent context
8 (akin to a 'classical stimulation' paradigm, see Hall and Plack, 2009) or a
9 noise context (akin to the 'continuous stimulation' paradigm, see Krumbholz et
10 al., 2003). For the silent context, we assume that the onset of each pitch will
11 evoke both energy- and pitch-onset responses. For the noise context, we
12 assume that the onset of each pitch will evoke only a pitch-onset response.
13 We assume that sustained responses to pitch (see Gutschalk et al 2004,
14 2007) are not markedly affected by the stimulus context and so do not
15 contribute to any observed differences in pitch-related activity. To ensure the
16 findings were not specific to a particular stimulus, the hypothesis was
17 examined using two different pitch-evoking stimuli; an unresolved harmonic
18 complex tone (UNRES) and a complex HP (cHP).

19

20 **Materials and methods**

21 *Listeners*

22 Fifteen listeners (8 male, 7 female; age range 23 – 48 years) with normal
23 hearing (≤ 20 dB hearing level between 250 Hz and 8 kHz) took part in this
24 study. All but one listener (#01) was right-handed (laterality index = 50,
25 Oldfield, 1971). Seven listeners were musically trained between grade 3 and

1 diploma level (# 01, 02, 07, 08, 10, 12 and 15) while five others reported
2 informal musical experience (self-taught/ungraded, # 04, 05, 09, 13 and 16).
3 One listener (#11) completed the psychophysical testing but was not able to
4 return for the fMRI session. None had a history of any neurological or hearing
5 impairment. Listeners gave written informed consent and the study was
6 approved by the Medical School Research Ethics Committee, University of
7 Nottingham.

8

9 *Stimuli*

10 All stimuli evoked a pitch corresponding to a 200-Hz tone. One stimulus was a
11 (diotic) unresolved-harmonic complex tone with harmonics 10-20 (henceforth
12 referred to as UNRES). The level of each harmonic was 23 dB greater than
13 the spectrum level of the **control** noise so that the gross spectral density of all
14 the stimuli was the same. The UNRES stimulus had a bandwidth of 2-4 kHz
15 with a noise masker (**49 dB SPL spectrum level**) from 0-2 kHz (to mask
16 cochlear distortion products). The other stimulus was a (dichotic) complex HP
17 (henceforth referred to as cHP) in which the pitch cue was only available via
18 integration of the signals from each ear (dichotic). The cHP stimulus was
19 created from a diotic Gaussian noise (**49 dB SPL spectrum level**) with a
20 bandwidth from 0-4 kHz. In one ear, a π phase shift was introduced in eight
21 30-Hz wide frequency bands, centered on the first eight harmonics. The noise
22 control stimulus was a Gaussian noise (**49 dB SPL spectrum level**), again low-
23 pass filtered at 4 kHz. The three signals (UNRES, cHP and noise) were
24 matched in bandwidth (0-4 kHz) and spectral density (and hence overall
25 energy). **It is probably impossible to generate stimuli that differ in pitch**

1 strength but are perfectly matched for every other perceptual feature. While
2 the Gaussian noise is a good control for cHP (the only other perceptual
3 difference between the two is the spatiality of the decorrelated band), it is
4 perhaps less so for UNRES because the signals differ in other respects, such
5 as their envelope structure. The logic of 'common activity' has been applied to
6 reduce the risk of attributing these potential differences to pitch (Hall and
7 Plack, 2009; Puschmann et al., 2010) and can also be applied to the analysis
8 of the current experiment.

9 For the psychophysical testing, each pitch stimulus was 350 ms in
10 duration (including 10-ms linear-intensity onset and offset ramps) and the
11 inter-stimulus interval was 500 ms. Reference stimuli had an f_0 of 200 Hz. The
12 stimuli were presented at an overall level of 85 dB SPL, calibrated using a
13 KEMAR manikin (Burkhard and Sachs, 1975) fitted with Bruel and Kjaer half-
14 inch microphone type 4134 (serial no. 906663), Zwislocki occluded ear
15 simulator (Knowles model no. DB-100) and Bruel and Kjaer measuring
16 amplifier type 2636 (serial no. 1324093), scaled from 22.4-Hz to 22.4 kHz
17 using fast time constant (125 ms) on maximum hold. Due to the metallic
18 components in the KEMAR system, calibration inside the scanner was not
19 possible.

20 In the scanner, stimulus conditions each comprised a 15.41-s
21 alternating sequence of 450-ms experimental sounds each separated by 230
22 ms. In the 'pitch-in-noise-context' conditions, the separation contained a
23 Gaussian noise as the context. In the 'pitch-in-silent-context' conditions, the
24 pitch signals were separated by 230 ms silence. The first and last components
25 of each sequence were the context. Each pitch and noise signal was

1 generated using 10 ms linear-intensity onset and offset ramps, which were
2 overlapped at the 3 dB points to produce a stable envelope for the stimulus
3 (see Figure 2). Eighteen sample sequences were created for each condition.

4 **The control noise conditions were created in the same way.**

5 ** INSERT FIGURE 2 AROUND HERE **

6 *Psychophysical testing*

7 Prior to the scanning session, each participant performed a pitch-
8 discrimination test to assess accuracy in distinguishing the pitch cues.
9 Psychophysical testing was carried out in a sound-attenuating booth and
10 stimuli were delivered through Sennheiser HD 480 II headphones. Stimuli
11 were presented through custom-made software that is supported by the
12 Matlab platform (The MathWorks, Natick, MA). Pitch discrimination thresholds
13 were measured for cHP and UNRES using a three alternative forced-choice,
14 two-down, one-up, adaptive procedure that targeted 70.7% performance
15 (Levitt, 1971). **Two observation intervals contained the standard tone (UNRES**
16 **or cHP) with an f₀ of 200 Hz. The remaining interval (chosen at random)**
17 **contained a comparison tone with a higher f₀ which the listener was required**
18 **to select as the 'odd one out'. On the first trial, the f₀ difference between**
19 **standard and comparison** was 20% (40 Hz). The percent difference increased
20 or decreased by a factor of two for the first four reversals, and by a factor of
21 1.414 for the final 12 reversals. Discrimination threshold was taken as the
22 geometric mean of the f₀ difference at the final 12 reversals. The adaptive
23 track was not allowed to increase above 200% (600 Hz). Responses were
24 recorded and stored electronically. On each trial, feedback was given via a
25 green (correct) or red (incorrect) light on the software interface. There were

1 five runs each for cHP and UNRES; the first was considered as practice and
2 so the pitch discrimination threshold was taken as the average of the last four
3 runs.

4

5 *fMRI protocol*

6 Scanning was performed on a Philips 3 T Intera Acheiva using an 8-channel
7 SENSE receiver head coil. A $T1$ -weighted high-resolution (1mm^3) anatomical
8 image (*matrix size* = 256×256 , 160 sagittal slices, $TR = 8.2$ ms, $TE = 3.7$ ms)
9 was collected for each subject. The anatomical scan was used to position the
10 functional scan centrally on HG, and care was taken to include the entire
11 superior temporal gyrus and to exclude the eyes. Functional scanning used a
12 $T2^*$ -weighted echo-planar sequence with a voxel size of 3mm^3 (*matrix size* =
13 64×64 , 32 oblique-axial slices, $TE = 36$ ms). Sparse imaging with a TR of
14 8000 ms and a clustered acquisition time of 1969 ms was used (Edmister et
15 al., 1999; Hall et al., 1999). A SENSE factor of 2 was applied to reduce image
16 distortions and a SofTone factor of 2 was used to reduce the background
17 scanner noise level by 9 dB. Functional data were acquired over two runs of
18 128 scans each, with the sounds presented in a quasi-random order, and with
19 the rule that the same stimulus condition was not presented twice in
20 succession. Listeners were requested to listen to the sounds, but were not
21 required to perform any task. A custom-built MR compatible system delivered
22 distortion-free sound using high-quality electrostatic headphones (Sennheiser
23 HE60 with high-voltage amplifier HEV70) that had been specifically modified
24 for use during fMRI.

25

1 *Data analysis*

2 Images were analyzed separately for each listener using statistical parametric
3 mapping (SPM5, <http://www.fil.ion.ucl.ac.uk/spm>). Preprocessing steps
4 included realignment to correct for subject motion, normalization of individual
5 scans to a standard image template, and smoothing with a Gaussian filter of 8
6 mm full width at half maximum. Individual analyses were computed for the two
7 runs (256 scans), specifying the two pitch and the two noise conditions as
8 separate regressors in the design. In the individual analysis, we specified
9 separate statistical contrasts for each sound condition relative to the silent
10 baseline that was implicitly modeled in the design. A high-pass filter cutoff of
11 420 s was used.

12 First, the data for individual participants was analyzed using a first-level
13 general linear model to assess the effects of interest with respect to the scan-
14 to-scan variability. The resulting model estimated the fit of the design matrix
15 (X) to the data (Y) in each voxel in order to provide β values (the contribution
16 of a single regressor to the overall fMRI signal). In order to obtain activation
17 maps for individual analysis, SPM was used to fit the GLM to each individual
18 voxel in the functional image, and to compute individual *t* statistics. The effect
19 of each stimulus condition was identified and the resulting (unthresholded)
20 contrast images were entered into a group-level random effects analysis in
21 order to assess the effects of interest with respect to the inter-subject
22 variability. At this group level, 2x2 repeated measures ANOVA were created,
23 with signal (pitch present and pitch absent, i.e. noise) and context (noise and
24 silent contexts) as factors. Separate ANOVA were computed for UNRES and
25 cHP conditions and within each ANOVA, simple main effects and interactions

1 were calculated using t statistics (Friston et al., 2005). Although initial SPM t
2 contrasts were defined using an uncorrected threshold of $p < 0.001$, all results
3 are reported after small volume correction (SVC) to control for type I errors
4 using a false discovery rate (FDR) threshold of $p < 0.05$ (Genovese et al 2002).
5 The small volume defined the auditory cortex across the superior temporal
6 gyrus (including HG, PT and planum polare) and contained 4719 voxels in the
7 left hemisphere and 5983 voxels in the right hemisphere. Activations were
8 localized using an SPM toolbox that overlays an SPM thresholded map onto a
9 set of probabilistic maps of the three cytoarchitectonic subdivisions of HG (Te
10 1.0, Te 1.1 and Te 1.2 Morosan et al., 2001; Eickhoff et al., 2005).

11

12 **Results**

13 *Behavioral results*

14 The mean geometric discrimination threshold across the listeners for cHP was
15 2.93 Hz and for UNRES was 3.54 Hz. A paired t -test showed that thresholds
16 for the two stimuli did not differ significantly [$t(1,14)=0.053$, $p=0.821$]. **For**
17 **comparison, previous research suggests that the threshold for a 200-Hz pure**
18 **tone is ~1 Hz (Wier et al., 1977), and that for a 200-Hz unresolved harmonic**
19 **complex tone is ~5 Hz (Houtsma and Smurzynski, 1990).**

20 ** INSERT FIGURE 3 AROUND HERE **

21 *fMRI results: Effect of stimulus context*

22 For the main effect of context, bilateral clusters of activation ($p < 0.05$ SVC)
23 were revealed for both pitch types (UNRES and cHP, Figure 3). These are
24 regions in which activation was greater for the silent context than for the noise
25 context. Both UNRES and cHP showed the same pattern of context-related

1 activation. In the left hemisphere, the most significant activation was in the
2 medial portion of HG (Te 1.1, Morosan et al., 2001) (x -38 y -26 z 6 mm for
3 both pitch types, Table 1). In the right hemisphere, the most significant
4 activation was in PT (cHP: x 64 y -30 z 12 mm; UNRES: x 66 y -22 z 10 mm).
5 However, for both pitch types there was substantial spread of context-related
6 activation across bilateral HG (areas Te 1.0, central HG; Te 1.1, medial HG,
7 and Te 1.2, lateral HG) and PT (Figure 3). There were no voxels that showed
8 a greater response for the noise context than the silent context, which
9 suggests that the human auditory cortex is more responsive to successive
10 energy onsets than it is to the overall energy in the stimulus.

11 ** INSERT TABLE 1 AROUND HERE **

12 *fMRI results: Effect of pitch*

13 The effect of signal (pitch present versus pitch absent) was also computed
14 from the 2x2 full factorial ANOVA to identify auditory cortical regions in which
15 activation was greater for the pitch condition than for the spectrally matched
16 noise control. Both cHP and UNRES contrasts revealed large bilateral clusters
17 of pitch-related activation (Table 1). For both pitch types, the greatest
18 response was located bilaterally in PT. Peak voxels were within 4 mm of each
19 other (x -58 y -24 z 8 mm and x -62 y -24 z 8 mm in the left hemisphere for
20 cHP and UNRES respectively; and again x 64 y -16 z 6 mm and x 66 y -18 z 6
21 mm in the right). While it is true that the pitch-related activity for UNRES
22 appeared somewhat more widespread than that for cHP, there was an
23 extremely high agreement between the most significant peaks (Table 1). This
24 finding increases our confidence that this activation focus represents a
25 response to the pitch quality of the stimuli, and not to some other feature that

1 was not perfectly matched between conditions. There were no voxels that
2 showed a significantly greater response to the control noise than to the pitch
3 stimulus.

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

22 ** INSERT FIGURE 4 AROUND HERE **

23 *Regional differences in the response to context and pitch*

24 The activation maps generated from the SPM analysis indicated a preference
25 for stimulus context in HG and a preference for pitch in PT and lack of any

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

14 ** INSERT FIGURE 5 AROUND HERE **

15 Figure 5c also demonstrates how PT is sensitive to UNRES and cHP.
16 For both contrasts, the main effect of pitch within PT reached significance
17 [$F(1,26)=38.76$, $p<0.001$ for UNRES and $F(1,26)=10.03$, $p=0.004$ for cHP]. In
18 Te 1.2, there was support for a preference for UNRES compared to the
19 control noise [$F(1,26)=9.72$, $p=0.004$], but this was not true for cHP
20 ($p=0.146$)(see Figure 5b). To examine whether these apparent regional
21 differences were significant, we performed a direct comparison between Te
22 1.2 and PT, again using ANOVA statistics. The overall effect of pitch was
23 significant [$F(1,26)=18.82$, $p<0.001$ for UNRES and $F(1,26)=4.58$, $p=0.042$ for
24 cHP] but there were no significant interactions between pitch and region for
25 either stimulus type. In other words, the effect of pitch was not significantly

1 greater in PT than in Te 1.2. It should be noted that the estimates of pitch-
2 related activity in PT are rather conservative because the estimates of
3 response magnitude were averaged over a large number of voxels, and many
4 of those in posterior PT were unresponsive to any of the stimulus features.
5 Thus, taken together, these analyses indicate a clear pitch-related response in
6 PT, but they do not rule out the possibility of a similar pitch response in Te 1.2
7 (i.e., lateral HG).

8

9 *Interaction showing modulatory effect of context on pitch-related activity*

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

20 To understand the shape of the interaction, a post-hoc region-of-
21 interest analysis was computed again using the estimates of size of the
22 response in PT. As Figure 5c shows, although the responses in the noise
23 context were marginally smaller than in the silent context, the difference
24 between the pitch and noise conditions was much more marked in the noise

1 context. Post-hoc testing demonstrated that this difference reached
2 significance ($p < 0.05$) in the UNRES condition.

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

12

13 **Discussion**

14 The present fMRI study reports a novel attempt to measure the effects of
15 stimulus context on the cortical representation of pitch. Concurring with
16 previous findings from the same authors (Hall and Plack, 2009), the most
17 significant pitch-related activity was centered on posterior auditory cortex, in
18 lateral PT. The results are consistent with the view that posterior auditory
19 cortex is engaged in nonspatial, as well as in spatial, auditory analysis (Arnott
20 et al., 2004). Some pitch-related activity was identified in lateral HG (Te 1.2).
21 However, there was no convincing evidence for a general sensitivity to pitch in
22 this region because the effect was only significant for UNRES. The UNRES
23 contrast is not ideally controlled for non-pitch features such as temporal
24 modulations in the UNRES signal that are not present in the control noise.

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

21 Although the response to energy onsets was greatest within primary
22 auditory cortex (including Te 1.0, Figure 5a), this effect was by no means
23 restricted to primary regions. Given the widespread effects of stimulus
24 context, it is cautionary to note that previous fMRI reports of pitch-related
25 activity could have perhaps underestimated the cortical representations of

1 pitch whenever those experiments utilized the ‘classical stimulation’ paradigm.
2 For example, we speculate that this explanation might account for the rather
3 low consistency of pitch-related activity reported for the 16 listeners in Hall
4 and Plack (2009).

5

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

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

1 With respect to the present localization of the energy-onset response in
2 human primary auditory cortex, it is perhaps worthwhile returning to the
3 findings from the depth electrode study reported by Schönwiesner and Zatorre
4 (2008). Recall, they reported energy-onset responses recorded at an
5 electrode placed on medial HG and pitch-onset responses at an electrode
6 placed on lateral HG, suggesting spatially separate neural populations
7 responsive to the two sound attributes. Their results are somewhat consistent
8 with the hierarchical model that we propose to account for our fMRI data.
9 Certainly, we would agree that medial HG was strongly responsive to energy
10 onsets because this was the main activation site where the fMRI response
11 was significantly diminished when energy onsets were eliminated by the noise
12 context. In the present study, there was partial evidence for the engagement
13 of lateral HG in the response to pitch. However, the present fMRI data would
14 predict that if one is to place an electrode array **intracranially** across the
15 surface of PT instead of across HG, then an even greater pitch response
16 might be recorded.

17 It is generally accepted that ‘higher’ cortical regions encode abstract
18 qualities, such as spatial location or speech sound identity (Davis and
19 Johnsruide, 2003; Nelken, 2008; Rauschecker et al., 1995; Rauschecker,
20 1998). The human nonprimary region PT would fit into this conceptualization.
21 However, we do not claim that the role of primary auditory cortex is merely to
22 encode simple (i.e. physical) attributes of a stimulus. In fact, it would perhaps
23 be rather naïve to do so. For example, electrophysiological recordings have
24 established that primary auditory cortical neurons have complex response
25 properties, showing sensitivity to both low-level and high-level features of

1 sounds (Nelken, 2008; Kelly et al., 1988). Nevertheless, the present study
2 demonstrates a reasonably clear division between HG and PT in terms of
3 representations of context and pitch, respectively. Moreover, only subdivisions
4 of PT (the 'higher' cortical region) were sensitive to response interactions
5 between stimulus context and pitch representations.

6

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

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

17 To ensure the findings were not specific to a particular stimulus, the
18 effect of context on pitch representations was examined using two different
19 pitch-evoking stimuli. The finding that the location of responses to UNRES
20 and cHP are broadly comparable suggests that the pitch responsive region
21 observed in the present study is not stimulus specific. However, UNRES
22 elicited activation in a greater number of voxels, Although inter-listener
23 consistency was low in the Hall and Plack (2009) study, an incidence map of
24 their 16 individual listeners also demonstrates similar patterns for their
25 unresolved harmonic complex and simple Huggins pitch. Information on the

1 number of activated voxels is not available for that study, but they did not find
2 a significant difference in percentage signal change for the two conditions. In
3 the present study, most of the pitch-related activity was centered on PT in
4 accordance with Hall and Plack's (2009) findings. Future investigations using
5 a greater number of different pitch-evoking stimuli would be required to
6 determine whether the results reported here are general to all pitch stimuli or
7 are specific to certain types of pitch stimuli.

8

9 **Acknowledgements**

10 This study was supported by a PhD studentship awarded by the MRC Institute
11 of Hearing Research and MR scanning was paid for through MRC
12 infrastructure funding awarded to the same organization. The authors would
13 also like to thank Deafness Research UK who provided a travel grant to
14 present this study at the 2009 ARO MidWinter Meeting in Baltimore, MA.

15

16 **References**

17

- 18 [Altmann, C.F., Henning, M., Döring, M.K., Kaiser, J., 2008. Effects of feature-](#)
19 [selective attention on auditory pattern and location processing.](#)
20 [NeuroImage 41, 69–79.](#)
- 21 [Arnott, S.R., Binns, M.A., Grady, C.L., Alain, C., 2004. Assessing the auditory](#)
22 [dual-pathway model in humans. NeuroImage 22, 401-408.](#)
- 23 [Barrett, D.J.K., Hall, D.A., 2006. Response preferences for “what” and “where”](#)
24 [in human non-primary auditory cortex. NeuroImage 32, 968-977.](#)

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

3 Burkard, M.D., Sachs, R.M., 1975. Anthropometric manikin for acoustic
4 research. *Journal of the Acoustical Society of America* 58, 214-222.

5 Carlyon, R.P., Demany, L., Semal, C., 1992. Detection of across-frequency
6 differences in fundamental frequency. *Journal of the Acoustical Society*
7 *of America* 91, 279-292.

8 Chait, M., Poeppel, D., Simon, J.Z., 2006. Neural response correlates of
9 detection of monaurally and binaurally created pitches in humans.
10 *Cerebral Cortex* 16, 835-848.

11 Chatterjee, M., Peng, S-C., 2008. Processing F0 with cochlear implants:
12 Modulation frequency discrimination and speech intonation recognition.
13 *Hearing Research* 235, 143-156.

14 Cramer, E.M., Huggins, W.H., 1958. Creation of pitch through binaural
15 interaction. *Journal of the Acoustical Society of America* 30, 413-417.

16 Dale, A.M., Buckner, R.L., 1997. Selective averaging of rapidly presented
17 individual trials using fMRI. *Human Brain Mapping* 5, 329-340.

18 Darwin, C.J., Ciocca, V., Sandell, G.J., 1994. Effects of frequency and
19 amplitude-modulation on the pitch of a complex tone with a mistuned
20 harmonic. *Journal of the Acoustical Society of America* 95, 2631-2636.

21 Davis, M.H., Johnsrude, I.S., 2003. Hierarchical processing in spoken
22 language comprehension. *Journal of Neuroscience* 23, 3423-3431.

23 Devor, A., Dunn, A.K., Andermann, M.L., Ulbert, I., Boas, D.A., Dale, A.M.,
24 2003. Coupling of total hemoglobin concentration, oxygenation, and
25 neural activity in rat somatosensory cortex. *Neuron* 39, 353-359.

1 Edmister, W.B., Talavage, T.M., Ledden, P.J., Weiskoff, R.M., 1999. Improved
2 auditory cortex imaging using clustered volume acquisitions. *Human*
3 *Brain Mapping* 7, 89-97.

4 Eickhoff, S.B., Amunts, K., Mohlberg, H., Zilles, K., 2006. The human parietal
5 operculum. II. Stereotaxic maps and correlation with functional imaging
6 results. *Cerebral Cortex* 16, 268-279.

7 Friston, K.J., Price, C.J., Fletcher, P., Moore, C., Frackowiak, R.S.J., Dolan,
8 R.J., 1996. The trouble with cognitive subtraction. *NeuroImage* 4, 97-
9 104.

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

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

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

17 Griffiths, T.D., Büchel, C., Frackowiak, R.S.J., Patterson, R.D., 1998. Analysis
18 of temporal structure in sound by the human brain. *Nature*
19 *Neuroscience* 1, 422-427.

20 Gutschalk, A., Patterson, R.D., Rupp, A., Uppenkamp, S., Scherg, M., 2002.
21 Sustained magnetic fields reveal separate sites for sound level and
22 temporal regularity in human auditory cortex. *NeuroImage* 15, 207-216.

23 Gutschalk, A., Patterson, R.D., Scherg, M., Uppenkamp, S., Rupp, A., 2004.
24 Temporal dynamics of pitch in human auditory cortex. *NeuroImage* 22,
25 755-766.

- 1 Gutschalk, A., Patterson, R.D., Scherg, M., Uppenkamp, S., Rupp, A., 2007.
2 The effect of temporal context on the sustained pitch response in
3 human auditory cortex. *Cerebral Cortex* 17, 552-561.
- 4 Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q.,
5 Elliott, M.R. et al., 1999. "Sparse" temporal sampling in auditory fMRI.
6 *Human Brain Mapping* 7, 213-223.
- 7 Hall, D.A., Barrett, D.J.K., Akeroyd, M.A., Summerfield, A.Q., 2005. Cortical
8 representations of temporal structure in sound. *Journal of*
9 *Neurophysiology* 94, 3181-3191.
- 10 Hall, D.A., Edmondson-Jones, A.M., Fridriksson, J., 2006. Frequency and
11 periodicity coding in human auditory cortex. *European Journal of*
12 *Neuroscience* 24, 3601–3610.
- 13 Hall, D.A., Plack, C.J., 2009. Pitch processing sites in the human auditory
14 brain. *Cerebral Cortex* 19, 576-585.
- 15 Harms, M.P., Melcher, J.R., 2002. Sound repetition rate in the human auditory
16 pathway: Representations in the waveshape and amplitude of fMRI
17 activation. *Journal of Neurophysiology* 88, 1433-1450.
- 18 Houtsma, A.J.M., Smurzynski, J., 1990. Pitch identification and discrimination
19 for complex tones with many harmonics. *Journal of the Acoustical*
20 *Society of America* 87, 304-310.
- 21 Kaas, J.H., Hackett, T.A., 1998. Subdivisions of auditory cortex and levels of
22 processing in primates. *Audiology and Neurootology* 3, 73–85.
- 23 Kelly, J.B., Sally, S.L., 1988. Organization of auditory cortex in the albino rat:
24 Binaural response properties. *Journal of Neurophysiology* 59, 1756-
25 1769.

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

4 Krumbholz, K., Eickhoff, S.B., Fink, G.R., 2007. Feature- and object-based
5 attentional modulation in the human auditory 'where' pathway. *Journal*
6 *of Cognitive Neuroscience* 19, 1721–1733.

7 König, R., Sieluzycycki, C., Durka, P.J., 2007. Tiny signals from the human
8 brain: Acquisition and processing of biomagnetic fields in
9 magnetoencephalography. *Journal of Low Temperature Physics* 146,
10 697-718.

11 Konishi, S., Yoneyama, R., Itagaki, H., Uchida, I., Nakajima, K., Yato, H.,
12 Okajima, K., Koisumi, H., Mijashita, Y., 1996. Transient brain activity
13 used in magnetic resonance imaging to detect functional areas.
14 *Neuroreport* 8, 19-23.

15 Levitt, H., 1971. Transformed up-down methods in psychoacoustics. *Journal*
16 *of the Acoustical Society of America* 49, 467-477.

17 Mahaffey, R.B., 1967. Some effects of cosine amplitude modulation on pitch
18 of tone bursts. *Journal of the Acoustical Society of America* 41, 1593.

19 Melcher, J.R., Sigalovsky, I.S., Guinan Jr., J.J., Levine, R.A., 2000.
20 Lateralized tinnitus studied with functional magnetic resonance
21 imaging: Abnormal inferior colliculus activation. *Journal of*
22 *Neurophysiology* 83, 1058-1072.

23 Micheyl, C., Oxenham, A.J., 2004. Sequential F0 comparisons between
24 resolved and unresolved harmonics: No evidence for translation noise

1 between two pitch mechanisms. Journal of the Acoustical Society of
2 America 116, 3038-3050.

3 Moore, B.C.J., 2004. An introduction to the psychology of hearing (5th Edition).
4 Elsevier, London.

5 Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T.,
6 Zilles, K., 2001. Human primary auditory cortex: Cytoarchitectonic
7 subdivisions and mapping into a spatial references system.
8 NeuroImage 13, 684-701.

9 Nääätänen, R., Picton, T., 1987. The N1 wave of the human electric and
10 magnetic response to sound: A review and an analysis of the
11 component structure. Psychophysiology 24, 375-425.

12 Nelken, I., Bizley, J.K., Nodal, F.R., Ahmed, B., King, A.J., Schnupp, J.W.H.,
13 2008. Responses of auditory cortex to complex stimuli: Functional
14 organization revealed using intrinsic optical signals. Journal of
15 Neurophysiology 99, 1928–1941.

16 Nelken, I., 2008. Processing of complex sounds in the auditory system.
17 Current Opinion in Neurobiology 18, 413-417.

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

20 Paltoglou, .E., Sumner C.J., Hall, D.A. 209. Examining the role of frequency
21 specificity in the enhancement and suppression of human cortical
22 activity by auditory selective attention. Hearing Research, 257, 106–
23 118.

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

4 Penagos, H., Melcher, J.R., Oxenham, A.J., 2004. A neural representation of
5 pitch salience in nonprimary human auditory cortex revealed with
6 functional magnetic resonance imaging. Journal of Neuroscience 24,
7 6810-6815.

8 Plack, C.J., Oxenham, A.J., 2005. The psychophysics of pitch. In: Plack CJ,
9 Oxenham AJ, Fay RR, Popper AN. Pitch: Neural coding and
10 perception. Springer, New York.

11 Plomp, R., 1964. The ear as a frequency analyzer. Journal of the Acoustical
12 Society of America 36, 1628-1636.

13 Price, C.J., Moore, C.J., Friston, K.J., 1997. Subtractions, conjunctions, and
14 interactions in experimental design of activation studies. Human Brain
15 Mapping 5, 264-272.

16 Puschmann, S., Uppenkamp, S., Kollmeier, B., Thiel, C.M., 2010. Dichotic
17 pitch activates pitch processing centre in Heschl's gyrus. NeuroImage
18 49, 1641-1649.

19 Rauschecker, J.P., 1998. Cortical processing of complex sounds. Current
20 Opinion in Neurobiology 8, 516-521.

21 Rauschecker J.P., Tian, B., Hauser, M., 1995. Processing of complex sounds
22 in the Macaque nonprimary auditory cortex. Science 268, 111-114.

23 Schönwiesner, M., Zatorre, R.J., 2008. Depth electrode recordings show
24 double dissociation between pitch processing in lateral Heschl's gyrus

1 and sound onset processing in medial Heschl's gyrus. *Experimental*
2 *Brain Research* 187, 97-105.

3 Seither-Priesler, A., Krumbholz, K., Patterson, R., Seither, S., Lütkenhöner,
4 B., 2004. Interaction between the neuromagnetic responses to sound
5 energy onset and pitch onset suggests common generators. *European*
6 *Journal of Neuroscience* 19, 3073-3080.

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

9 Singh, P.G., 1987. Perceptual organization of complex-tone sequences: A
10 tradeoff between pitch and timbre? *Journal of the Acoustical Society of*
11 *America* 82, 886-899.

12 Westbury, C.F., Zatorre, R.J., Evans, A.C., 1999. Quantifying variability in the
13 planum temporale: A probability map. *Cerebral Cortex* 9, 392-405.

14 Wier, C., Jesteadt, W., Green, D., 1977. Frequency discrimination as a
15 function of frequency and sensation level. *Journal of the Acoustical*
16 *Society of America* 61, 178-184.

17 Yost, W.A., Hill, R., 1979. Models of the pitch and pitch strength of ripple
18 noise. *Journal of the Acoustical Society of America* 66, 400-410.

19 Yost, W., 1996. Pitch of iterated rippled noise. *Journal of the Acoustical*
20 *Society of America* 100, 511-518.

21

22 Figure legends

23

24 Figure 1. Two models depicting the coupling between neural activity and fMRI
25 activation. (a) The non-linear response model suggests that the fMRI

1 response is limited by a saturation level (dotted line) which, in the silent
2 context condition, is dominated by multiple energy onsets so that the addition
3 of a pitch elicits little additional activation. (b) The linear response model
4 suggests that the response is additive. In this case fMRI activation is identical
5 in silent and noise contexts.

6

7 Figure 2. Schematic diagram of the signal and context components of the
8 stimuli, overlapped to produce a stable envelope. In the 'pitch-in-noise-
9 context' and 'pitch-in-silent-context' conditions, the signal is either UNRES or
10 cHP and the context is Gaussian noise or silence, respectively. In the 'noise-
11 in-noise-context' condition, both signal and context segments are Gaussian
12 noise, hence it is a continuous noise. In the 'noise-in-silent-context' condition,
13 the signal is Gaussian noise and the context is silence.

14

15 Figure 3. Activation map from the 2x2 factorial ANOVA showing locations for
16 the main effects of context (cyan) and signal (magenta), regions where the
17 two main effects overlap (purple), and areas in which context modulates pitch
18 (yellow). The white borders denote areas Te 1.1 (medial portion), Te 1.0
19 (middle portion) and Te 1.2 (lateral portion) (Morosan et al., 2001) on Heschl's
20 gyrus. The black border outlines PT (Westbury et al., 1999). Activation is
21 overlaid onto an average anatomical image made from the 15 individual
22 listeners. **The left hemisphere is on the left-hand side of each anatomical**
23 **image.**

24

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

12

13 Figure 5. Graphical representation of the effects of context and pitch, and their
14 interactions plotted separately for UNRES and cHP. The three panels
15 represent the three different regions of interest: (a) primary auditory cortex
16 defined by Te 1.0, (b) lateral HG defined by Te 1.2, and (c) PT. Error bars
17 represent 95% confidence intervals around the estimated mean activation.

18

4. Table

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.

cHP contrasts

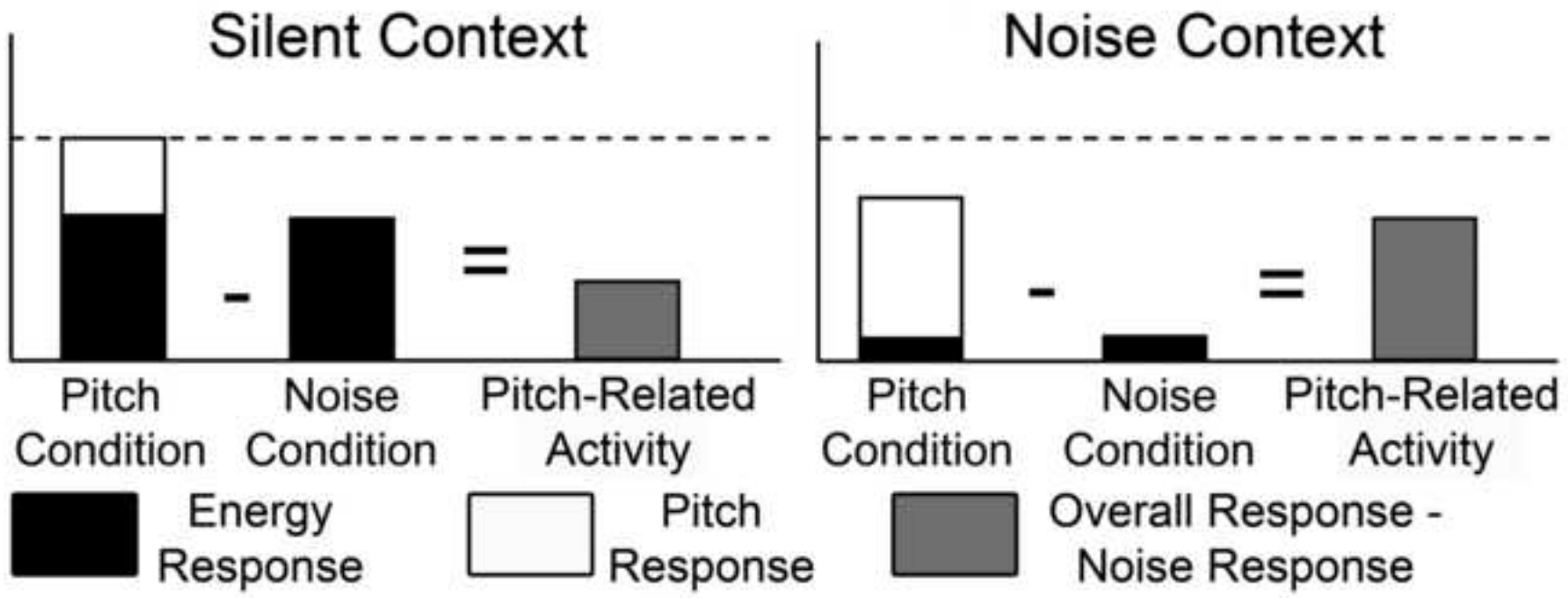
	Left Hemisphere					Right Hemisphere				
	Peak co-ordinates	Z-score	Voxel-level p-value*	Cluster size	Location	Peak co-ordinates	Z-score	Voxel-level p-value*	Cluster size	Location
Context: silence> noise	-38 -26 6	5.79	<0.001	1181	Te 1.1	64 -30 12	5.36	<0.001	1182	PT
	-66 -42 20	3.39	0.003	8	PT	-	-	-	-	-
	-42 -14 20	3.34	0.003	23	Te 1.0	-	-	-	-	-
	-44 -20 14	3.18	0.005	2	Te 1.0	-	-	-	-	-
Pitch:cHP >noise	-58 -24 8	4.48	0.002	314	PT	64 -16 6	5.5	<0.001	137	PT
	-48 -6 -8	3.22	0.019	3	PP	56 -36 4	3.09	0.023	1	PT
Interaction	No suprathreshold voxels					No suprathreshold voxels				

UNRES contrasts

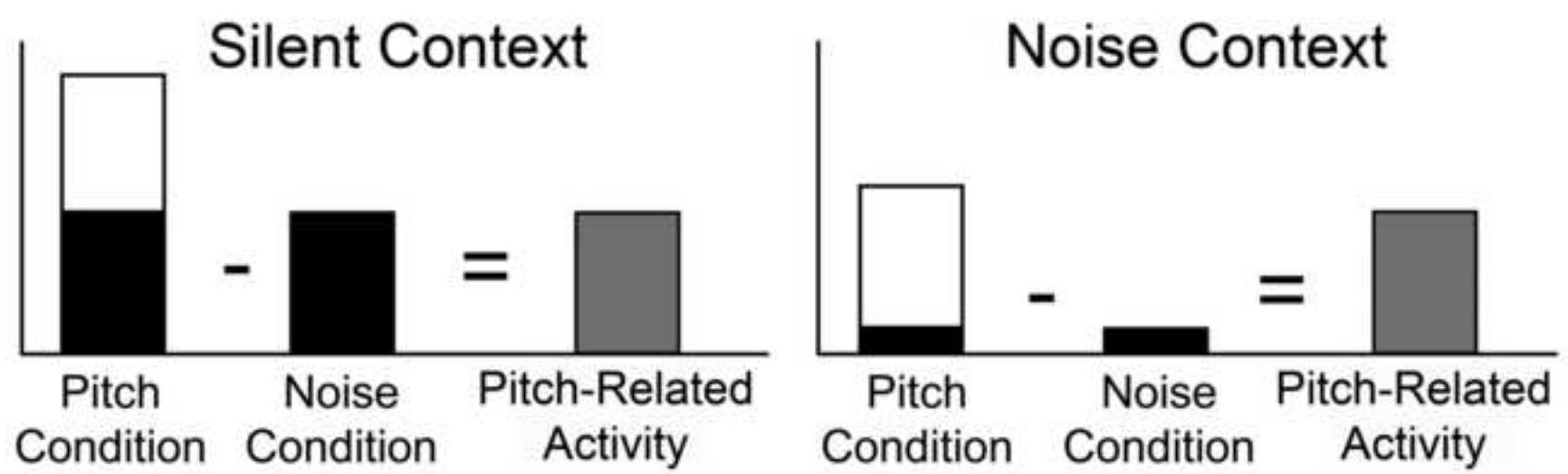
	Left Hemisphere					Right Hemisphere				
	Peak co-ordinates	Z-score	Voxel-level p-value*	Cluster size	Location	Peak co-ordinates	Z-score	Voxel-level p-value*	Cluster size	Location
Context: silence> noise	-38 -26 6	5.28	<0.001	1375	Te 1.1	66 -22 10	4.49	0.001	550	PT
	-	-	-	-	-	62 -4 8	4.01	0.002	29	Te 1.0
	-60 -38 14	3.43	0.004	24	PT	66 -36 24	3.35	0.005	5	PT
	-36 -18 16	3.26	0.006	2	Te 1.1	38 -20 -2	3.18	0.007		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	0.004	29	Te 1.0	62 -4 2	3.86	0.001	22	Te 1.2
	-60 -12 4	3.35	0.004	14	Te 1.2	46 -14 -8	3.21	0.005	4	PP
	-64 -40 20	3.56	0.04	20	PT	36 -30 4	4.19	0.038	45	PT
	-46 -18 -2	3.47	0.043	16	Te 1.0	68 -30 12	4.11	0.038	234	PT
Interaction	-44 -38 20	3.19	0.049	1	PT	42 -40 16	3.24	0.046	2	PT

*FDR-corrected

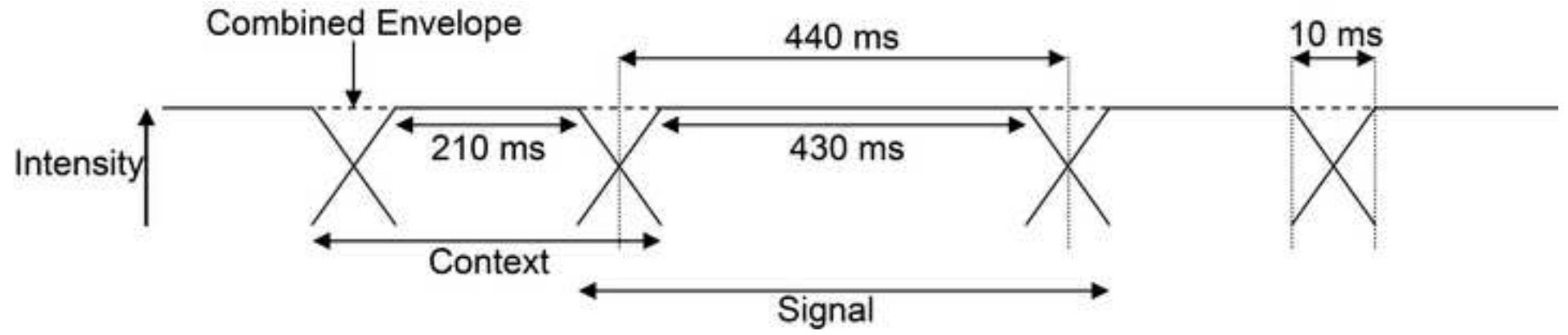
a. Non-Linear Response Model



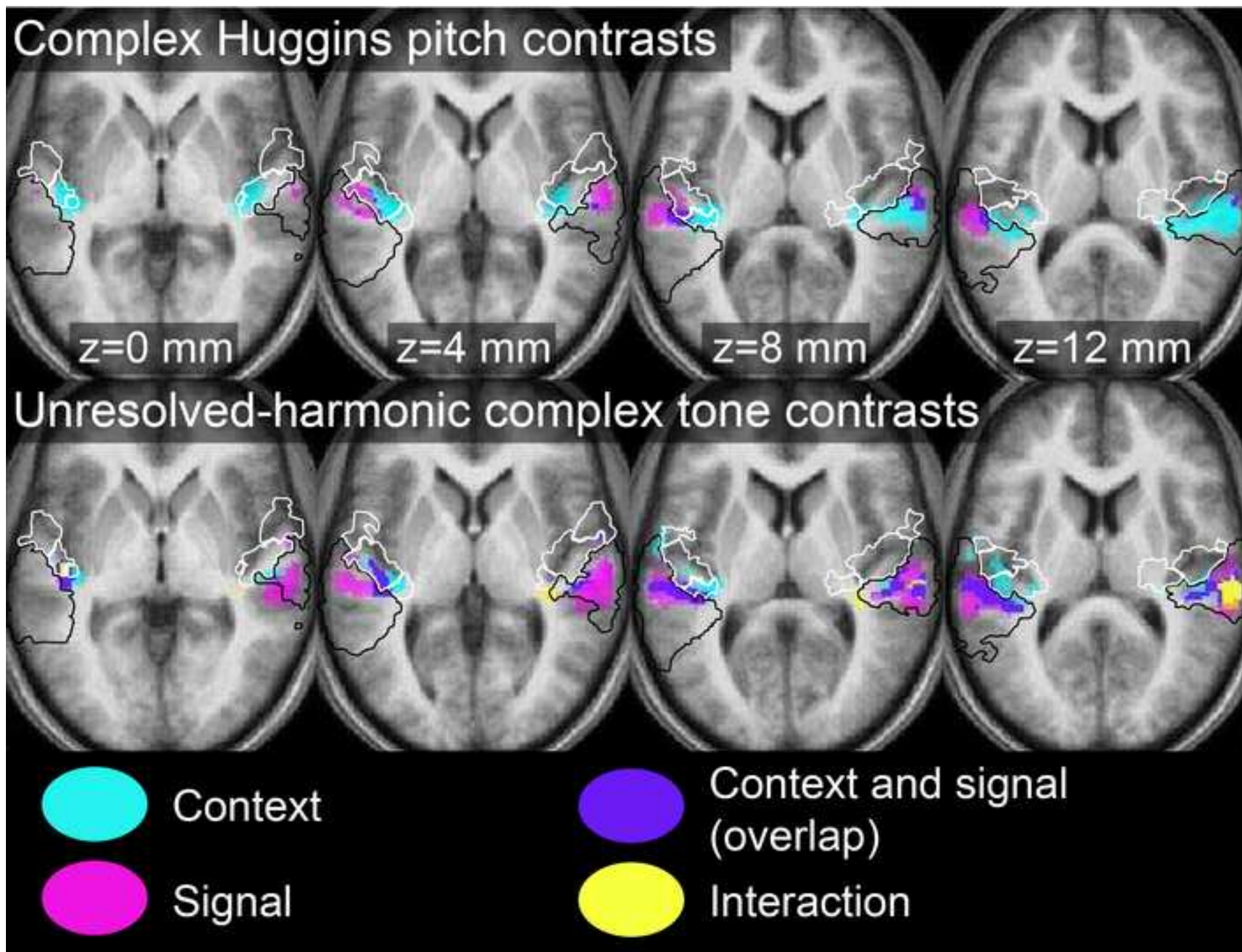
b. Linear Response Model



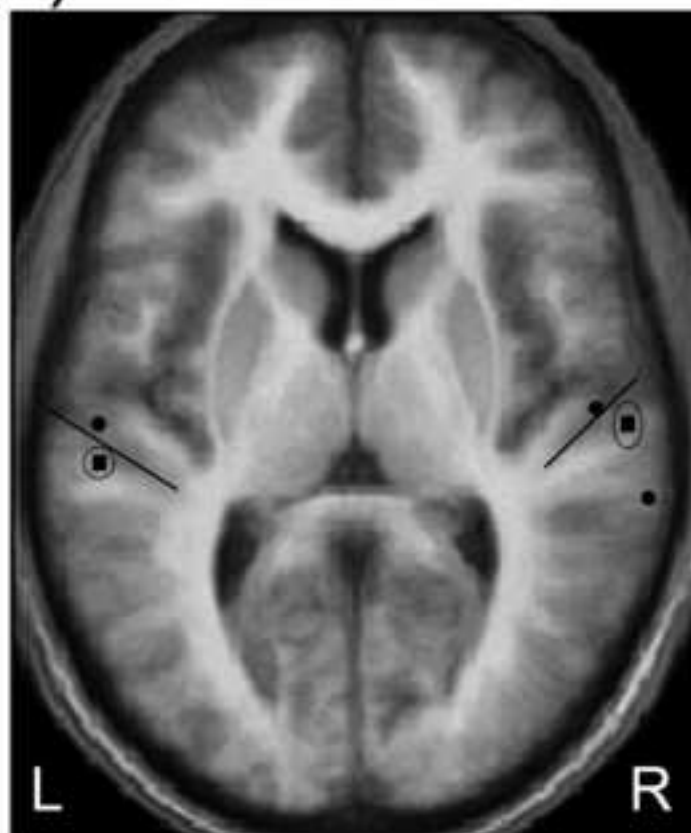
5. Figure
[Click here to download high resolution image](#)



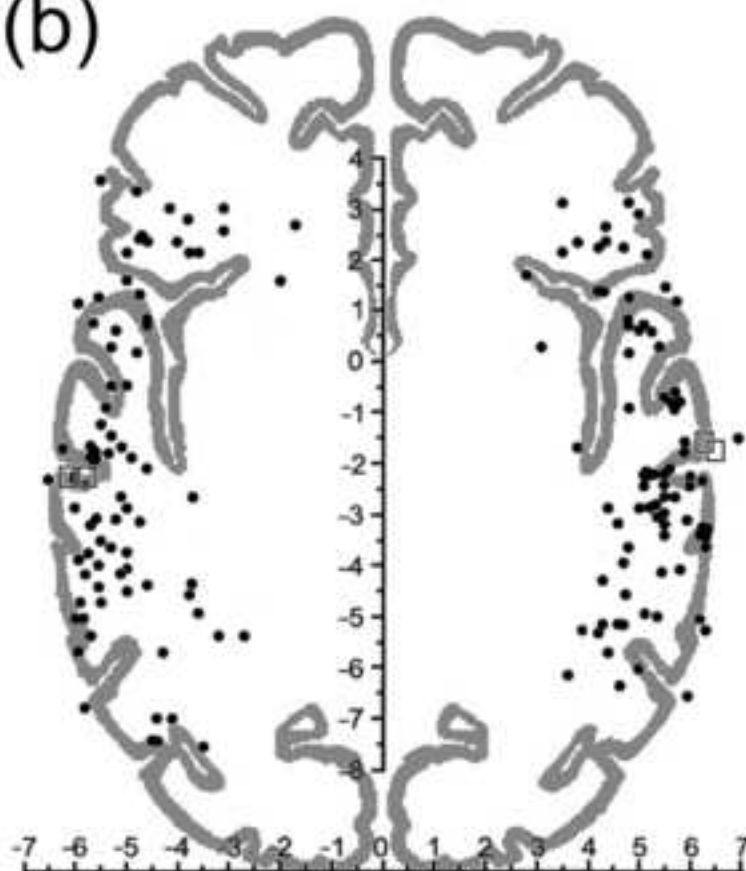
5. Figure
[Click here to download high resolution image](#)



(a)



(b)



5. Figure
[Click here to download high resolution image](#)

