

Mapping Human Pitch Representation in a Distributed System Using Depth-Electrode Recordings and Modeling

Sukhbinder Kumar^{1,2} and Marc Schönwiesner^{3,4}

¹Institute of Neuroscience, Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH, United Kingdom, ²Wellcome Trust Center for Neuroimaging, 12 Queen Square, London WC1N 3BG, United Kingdom, ³International Laboratory for Brain, Music and Sound Research, Montreal, Montreal, Quebec, H3C 3J7 Canada, and ⁴Department of Psychology, University of Montreal, Montreal, Quebec, H2V 2S9, Canada

Depth-electrode recordings from the auditory cortex of humans undergoing presurgical evaluation for epilepsy allow the recording of ensemble responses to pitch in the form of local field potentials. These recordings allow another test of the hypothesis that there is a specialized neural ensemble for pitch within auditory cortex. Moreover, the technique allows recordings from multiple sites with millisecond temporal resolution to allow modeling of the effective connectivity between these sites. Here we argue that this takes the form of a hierarchical network of pitch-sensitive regions. Activity can be understood as reflecting predictive coding, in which perceptual predictions and error messages are continuously exchanged between a higher pitch center and lower-level auditory cortex.

Evidence from intracranial recordings for a lateral pitch center

Human fMRI studies (Griffiths et al., 1998; Patterson et al., 2002; Penagos et al., 2004) and lesion studies (Zatorre, 1988; Johnsrude et al., 2000) suggest an area in the cerebral cortex that is specialized in the representation of pitch in non-primary auditory cortex, specifically in the vicinity of lateral Heschl's gyrus (HG) (see review by Griffiths and Hall, 2012). Other studies have also implicated this region in computational functions that seem crucial for extracting pitch from complex sounds, such as spectral information integration over time or frequency channels (Zatorre and Belin, 2001; Hall et al., 2002; Schönwiesner et al., 2005). The proposal of a "pitch center" in human lateral non-primary auditory cortex can be directly tested by recording intracerebral potentials using chronically implanted electrodes. This technique has high temporal and spatial resolution, making it possible to locate the cerebral sources of pitch-related responses and relate their temporal characteristics to previous results from noninvasive techniques with high temporal resolution, such as magneto- and electroencephalography. Only two previous studies have recorded intracerebral potentials related to pitch processing with depth electrodes (Schönwiesner and Zatorre, 2008; Griffiths et al., 2010). Here, we review these studies and provide a novel framework of distributed pitch processing. Results of these studies are schematically summarized in Figure 1. Schönwiesner and Zatorre (2008) measured responses from different locations along HG to transitions from a non-pitched stimulus to a spectrally matched but temporally regular noise (regular-interval noise, RIN) that evokes a percept of pitch (for more details of the

RIN stimulus, see Yost, 1996). The RIN was preceded by a 1-s-long noise without pitch so that the transient response evoked by the sound onset would subside before the pitch onset. This makes it possible to isolate pitch-related responses in time from other processes. Krumbholz et al. (2003) used this design in an MEG experiment to isolate a transient component of the auditory evoked field that indicates pitch processing ("pitch onset response"). Results from intracerebral recordings showed a pitch-onset response at an electrode contact close to lateral HG, but not at medial contacts. This lateral site was within non-primary auditory cortex, because responses to pure tones were less frequency-specific and had longer latencies than responses from primary auditory cortex at medial recording sites.

Evidence from intracranial recordings for distributed pitch responses

Recent studies have collected evidence for the involvement of other regions of auditory cortex in pitch processing in both humans and animals. Griffiths et al. (2010) recorded local field potentials in response to transition from noise to RIN at multiple contacts along the axis of HG in two human subjects. In one experiment, rate of periodicity of RIN was varied from 8 Hz to 256 Hz. Taking advantage of the fact that the percept of pitch is evoked only when the rate of periodicity is above ~30 Hz (lower limit of pitch, Krumbholz et al., 2000; see also review by Oxenham, 2012), Griffiths et al. (2010) dissociated brain responses that correlate with the stimulus feature of temporal regularity from the responses that reflect the pitch percept. They observed that phase-locked responses occurred at all rates of periodicities of RIN, both below and above the lower limit of pitch, whereas induced response in the high gamma range (80–120 Hz) occurred only when the rate of periodicity was above 32 Hz. This shows that while the evoked response may reflect the representation of stimulus features, the induced response in the gamma range may be a correlate of the pitch. Both evoked and induced

Received April 4, 2012; revised July 18, 2012; accepted July 23, 2012.

Correspondence should be addressed to either of the following: Sukhbinder Kumar, Institute of Neuroscience, Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH, UK, E-mail: sukhbinder.kumar@ncl.ac.uk, or Marc Schönwiesner at the above address, Department of Psychology, University of Montreal, Montreal, Quebec, H2V 2S9, Canada, E-mail: marc.schoenwiesner@umontreal.ca.

DOI:10.1523/JNEUROSCI.3812-12.2012

Copyright © 2012 the authors 0270-6474/12/3213348-04\$15.00/0

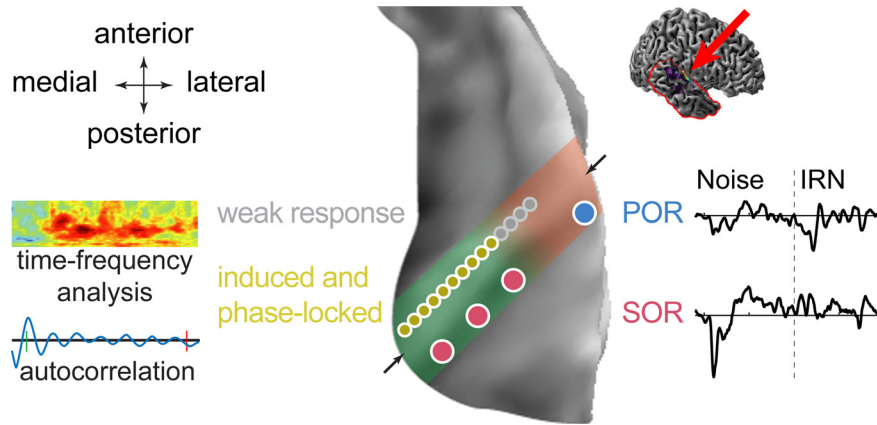


Figure 1. Summary of previous studies on pitch processing using depth-electrode recordings. Positions of electrode contacts are illustrated as colored circles on the surface of the right superior temporal plane (viewpoint shown with red arrow in the top-left inset). Heschl’s gyrus is between the black arrows, with primary auditory cortex highlighted in green and lateral, non-primary cortex highlighted in red. Griffiths et al. (2010) found induced and phase-locked responses in medial and middle Heschl’s gyrus (yellow circles) and weak responses at more lateral contacts (gray circles). Induced responses were found as gamma-band oscillations in a time-frequency analysis. Phase-locked responses were demonstrated by a significant peak in the autocorrelation of the evoked responses at the inverse of the pitch frequency. Schönwiesner and Zatorre (2008) found strong pitch-onset responses (POR), but only small sound-onset responses, at the most lateral contact (blue circle) to the onset of RIN pitch in an ongoing noise. Medial electrode contacts (red circles) showed strong sound-onset responses (SOR), but small or no pitch-onset responses.

responses were distributed along medial and middle HG, but were weaker at the more lateral contacts.

In another experiment, Griffiths et al. (2010) varied salience of pitch of RIN, with the pitch value kept fixed. As the pitch salience increased, the magnitude of both the evoked and induced responses increased at all locations along the HG. The distributed sensitivity to pitch salience is in apparent contradiction with results from Bendor and Wang (2010), where neurons that increased their firing rate with increase in pitch salience were located only in the pitch center.

Evidence from Griffiths et al. (2010) shows that pitch-related information is available in multiple areas of the auditory cortex. Further evidence for distributed representation of pitch also comes from studies in animals. In ferrets, Walker et al. (2011) found sensitivity to pitch in all auditory areas that were examined in the study, including primary auditory cortex (see also review by Wang and Walker, 2012).

Modeling pitch perception

Given the evidence that pitch-related responses are present in multiple areas of the auditory cortex, an immediate concern is how the different areas interact during pitch processing. Moreover, a single pitch center in auditory cortex is in contradiction with the distributed view. Could the two views be reconciled by considering that one area plays a dominant role, thereby incorporating the role of a pitch center within a distributed system? This hypothesis can be tested with a network-level analysis of the activity observed across multiple auditory areas.

One influential model of perception, the predictive coding model (Rao and Ballard, 1999; Friston and Kiebel, 2009), provides an explanation of how information is integrated across multiple areas during perception. This model has found strong empirical support both in the visual (Srinivasan et al., 1982; Dong and Atick, 1995; Hosoya et al., 2005; Jehee et al., 2006) and auditory (Smith and Lewicki, 2006; Vuust et al., 2009; for review, see Winkler et al., 2009; Winkler and Czigler, 2012) modalities. The predictive coding model of brain function posits that the brain holds an internal model of the world which is embedded in the cortical hierarchy of the brain. Using this internal model, areas at higher level of hierarchy actively predict the input they expect to receive from lower areas. This prediction is passed to the lower

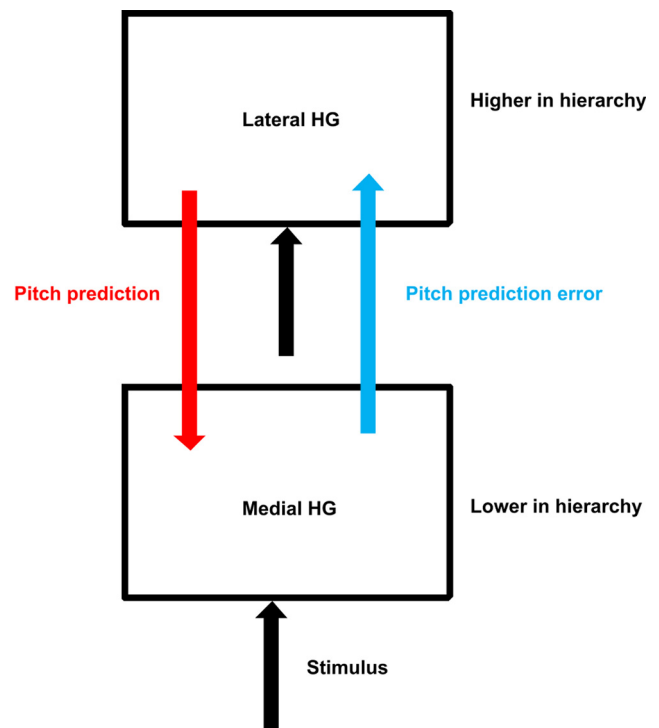


Figure 2. Schematic representation of predictive coding model of pitch perception. Both lateral and medial parts of Heschl’s gyrus receive stimulus independently from the auditory thalamus (black arrow). The lateral HG, makes a prediction of pitch which is passed to the medial HG via backward connections (red arrow), where prediction error (difference between the prediction received from lateral HG and stimulus representation at medial HG) is computed which is passed to the lateral HG by means of forward connections (blue arrow).

area via backward connections. The lower area computes prediction error (difference between the prediction received from the area above and the representation at that level), which is passed to the higher area via forward connections. The strength of forward and backward connections is constantly adjusted so as to minimize prediction error.

The predictive coding model of pitch perception entails that the auditory system actively predicts the pitch of an ongoing

stimulus rather than passively extracting the stimulus features from which the percept is generated. Evidence that the auditory system can predict the future course of the stimulus from the past temporal regularity of the stimulus at fine temporal resolution comes from a number of studies (Grimm and Schröger, 2007; Grimm et al., 2011; for review, see Bendixen et al., 2012). If pitch perception could be explained by the predictive coding, then this model makes the following specific predictions. (1) Areas of the pitch system should be organized hierarchically. This is because areas at different levels of hierarchy integrate stimulus information at different temporal scales (Okada et al., 2010). Lower-level areas may track local fluctuations in spectro-temporal features, but may not integrate over a time scale sufficient for producing a pitch percept, whereas higher-level areas may integrate information over a time scale suitable for producing a pitch percept. These higher-level areas may thus predict pitch value more reliably than lower-level areas. (2) Connection strength between hierarchical levels should change as a function of pitch saliency of the stimulus. This is because with a higher pitch saliency, pitch value can be predicted with greater reliability which in turn decreases prediction error computed at the lower level in hierarchy. Because prediction is conveyed by the top-down connections and prediction error is conveyed by the bottom-up connections, the strength of the top-down connection would increase, and, correspondingly, the strength of the bottom-up connection would decrease with greater pitch saliency.

Kumar et al. (2011) further analyzed the data reported by Griffiths et al. (2010) using dynamic causal modeling (DCM) (David et al., 2006). The central idea behind DCM is to identify causal interactions (effective connectivity) between two or more areas. The term “causal” in DCM refers to how the activity of one brain area changes the dynamics and/or response of another area. In addition to quantifying effective connectivity between areas, DCM also allows the comparison of hierarchical architectures within the auditory system by defining forward connections (from lower to higher areas), parallel connections (between areas at the same hierarchical level), and backward connections (from higher to lower areas; Felleman and Van Essen, 1991). Without prior assumption of the processing hierarchy of areas along HG, Kumar et al. (2011) generated an exhaustive set of all possible hierarchical configurations of medial, middle, and lateral HG. Using Bayesian model comparison to determine the configuration, or the model that explains the data best, it was shown that areas on the HG were arranged in the following hierarchical configuration: lateral part of HG is at a higher level of hierarchy than the medial and middle part of HG, both of which are at the same levels of hierarchy. This hierarchical configuration is in agreement with the evidence from depth-electrode recordings along HG, showing that the medial and middle electrode contacts are in primary auditory cortex, whereas the lateral contacts are in non-primary cortex (Brugge et al., 2009). Further analysis of how this connectivity varies with pitch saliency showed that strength of backward connection from lateral HG to both medial and middle HG increased with pitch saliency, whereas the strength of forward connections from medial and middle HG to lateral HG decreased with saliency in accordance with predictive coding mechanism of pitch perception. A schematic of the proposed predictive coding model of pitch perception is shown in Figure 2.

The localized and distributed views of pitch perception may not be completely antagonistic to each other. In the above model, we find that lateral HG plays more of a “dominant” role in pitch perception in the sense that it is involved in actively predicting the pitch of the stimulus. Lateral HG may thus comprise pitch-

specific processing mechanisms that lead increased responses to pitched stimuli compared with non-pitched stimuli in functional neuroimaging studies.

Conclusions and future directions

There are two features of the above model that are distinct from the models of pitch perception proposed previously (for review, see Cheveigné, 2005). First, in addition to the bottom-up flow of information, the pitch system has a top-down component. Second, the pitch system is hierarchical; that is, information is processed at multiple levels. Most previous models of pitch perception lack both of these features (for a counterexample, see Balaguer-Ballester et al., 2009), which makes these models less biologically plausible. Evidence from psychophysical studies also shows that a top-down component plays a role in pitch percept: whether a mistuned component of harmonic complex changes the pitch of a harmonic complex or not depends on the context in which it occurs (Darwin and Ciocca, 1992). Similarly, the evidence that the pitch system can use different time scales to compute the percept of pitch (Plack and Oxenham, 2005) shows the importance for the system to be hierarchical.

Limitations of the evidence and the model presented here need to be recognized. First, evidence from the human fMRI suggests that pitch-related information may be available in areas outside the HG. Because of the limited coverage of depth-electrode recordings, contributions of those areas to pitch perception have not been assessed. Second, the model does not distinguish between representations of pitch and of other features that covary with pitch, such as temporal regularity. The model is therefore equally valid for those acoustic features that correlate with pitch.

The animal work on pitch perception has been mainly informed by recordings from single or multiple units in the auditory cortex. In contrast, human neurophysiology work has focused on local field potentials. In the future, recording of multiunit activity from high-impedance electrodes (Howard et al., 1996) will allow studying human pitch processing at the single cellular level and potentially bridge the gap between animal and human results.

References

- Balaguer-Ballester E, Clark NR, Coath M, Krumbholz K, Denham SL (2009) Understanding pitch perception as a hierarchical process with top-down modulation. *PLoS Comput Biol* 5:e1000301. [CrossRef Medline](#)
- Bendixen A, SanMiguel I, Schröger E (2012) Early electrophysiological indicators for predictive processing in audition: a review. *Int J Psychophysiol* 83:120–131. [CrossRef Medline](#)
- Bendor D, Wang X (2005) The neuronal representation of pitch in primate auditory cortex. *Nature* 436:1161–1165. [CrossRef Medline](#)
- Bendor D, Wang X (2010) Neural coding of periodicity in marmoset auditory cortex. *J Neurophysiol* 103:1809–1822. [CrossRef Medline](#)
- Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA 3rd (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J Neurophysiol* 102:2358–2374. [CrossRef Medline](#)
- Cheveigné A (2005) Pitch perception models. In: *Pitch* (Plack C, Fay R, Oxenham A, Popper A, eds), pp 169–233. New York: Springer.
- Darwin CJ, Ciocca V (1992) Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component. *J Acoust Soc Am* 91:3381–3390. [CrossRef Medline](#)
- David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ (2006) Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage* 30:1255–1272. [CrossRef Medline](#)
- Dong DW, Atick JJ (1995) Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus. *Netw Comput Neural Sys* 6:159–178.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47. [CrossRef Medline](#)

- Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B Biol Sci* 364:1211–1221. [CrossRef Medline](#)
- Griffiths TD, Büchel C, Frackowiak RS, Patterson RD (1998) Analysis of temporal structure in sound by the human brain. *Nat Neurosci* 1:422–427. [CrossRef Medline](#)
- Griffiths TD, Kumar S, Sedley W, Nourski KV, Kawasaki H, Oya H, Patterson RD, Brugge JF, Howard MA (2010) Direct recordings of pitch responses from human auditory cortex. *Curr Biol* 20:1128–1132. [CrossRef Medline](#)
- Grimm S, Schröger E (2007) The processing of frequency deviations within sounds: evidence for the predictive nature of the Mismatch Negativity (MMN) system. *Restor Neurol Neurosci* 25:241–249. [Medline](#)
- Grimm S, Escera C, Slabu L, Costa-Faidella J (2011) Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology* 48:377–384. [CrossRef Medline](#)
- Hall DA, Johnsrude IS, Haggard MP, Palmer AR, Akeroyd MA, Summerfield AQ (2002) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 12:140–149. [CrossRef Medline](#)
- Hosoya T, Baccus SA, Meister M (2005) Dynamic predictive coding by the retina. *Nature* 436:71–77. [CrossRef Medline](#)
- Howard MA 3rd, Volkov IO, Granner MA, Damasio HM, Ollendieck MC, Bakken HE (1996) A hybrid clinical-research depth electrode for acute and chronic in-vivo microelectrode recording of human brain neurons. *J Neurosurg* 84:129–132. [CrossRef Medline](#)
- Jehee JFM, Rothkopf C, Beck JM, Ballard DH (2006) Learning receptive fields using predictive feedback. *J Physiol Paris* 100:125–132. [CrossRef Medline](#)
- Johnsrude IS, Penhune VB, Zatorre RJ (2000) Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123:155–163. [CrossRef Medline](#)
- Krumbholz K, Patterson RD, Pressnitzer D (2000) The lower limit of pitch as determined by rate discrimination. *J Acoust Soc Am* 108:1170–1180. [CrossRef Medline](#)
- Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhöner B (2003) Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb Cortex* 13:765–772. [CrossRef Medline](#)
- Kumar S, Sedley W, Nourski KV, Kawasaki H, Oya H, Patterson RD, Howard MA 3rd, Friston KJ, Griffiths TD (2011) Predictive coding and pitch processing in the auditory cortex. *J Cogn Neurosci* 23:3084–3094. [CrossRef Medline](#)
- Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences JT, Hickok G (2010) Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cereb Cortex* 20:2486–2495. [CrossRef Medline](#)
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36:767–776. [CrossRef Medline](#)
- Penagos H, Melcher JR, Oxenham AJ (2004) A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci* 24:6810–6815. [CrossRef Medline](#)
- Plack C, Oxenham AJ (2005) The psychophysics of pitch. In: *Pitch: neural coding and perception* (Plack CJ, Oxenham AJ, Fay RR, eds), pp 7–56. New York: Springer.
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87. [CrossRef Medline](#)
- Schönwiesner M, Zatorre RJ (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. [CrossRef Medline](#)
- Schönwiesner M, Rübsem R, von Cramon DY (2005) Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur J Neurosci* 22:1521–1528. [CrossRef Medline](#)
- Smith EC, Lewicki MS (2006) Efficient auditory coding. *Nature* 439:978–982. [CrossRef Medline](#)
- Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci* 216:427–459. [CrossRef Medline](#)
- Vuust P, Ostergaard L, Pallesen KJ, Bailey C, Roepstorff A (2009) Predictive coding of music—brain responses to rhythmic incongruity. *Cortex* 45:80–92. [CrossRef Medline](#)
- Walker KMM, Bizley JK, King AJ, Schnupp JWH (2011) Multiplexed and robust representations of sound features in auditory cortex. *J Neurosci* 31:14565–14576. [CrossRef Medline](#)
- Winkler I, Czigler I (2012) Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *Int J Psychophysiol* 83:132–143. [CrossRef Medline](#)
- Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn Neurosci* 13:532–540. [CrossRef](#)
- Yost WA (1996) Pitch of iterated rippled noise. *J Acoust Soc Am* 100:511–518. [CrossRef Medline](#)
- Zatorre RJ (1988) Pitch perception of complex tones and human temporal-lobe function. *J Acoust Soc Am* 84:566–572. [CrossRef Medline](#)
- Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11:946–953. [CrossRef Medline](#)