J Neurophysiol 107: 2042–2056, 2012.
First published February 1, 2012; doi:10.1152/jn.00308.2011.

Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fMRI evidence

Tobias Overath,[1] Yue Zhang,[1] Dan H. Sanes,[2,3] and David Poeppel[1,2]
[1]*Department of Psychology,* [2]*Center for Neural Science, and* [3]*Department of Biology, New York University, New York, New York*

Submitted 6 April 2011; accepted in final form 29 January 2012

Overath T, Zhang Y, Sanes DH, Poeppel D. Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fMRI evidence. *J Neurophysiol* 107: 2042–2056, 2012. First published February 1, 2012; doi:10.1152/jn.00308.2011.—Hierarchical models of auditory processing often posit that optimal stimuli, i.e., those eliciting a maximal neural response, will increase in bandwidth and decrease in modulation rate as one ascends the auditory neuraxis. Here, we tested how bandwidth and modulation rate interact at several loci along the human central auditory pathway using functional MRI in a cardiac-gated, sparse acquisition design. Participants listened passively to both narrowband (NB) and broadband (BB) carriers (1/4- or 4-octave pink noise), which were jittered about a mean sinusoidal amplitude modulation rate of 0, 3, 29, or 57 Hz. The jittering was introduced to minimize stimulus-specific adaptation. The results revealed a clear difference between spectral bandwidth and temporal modulation rate: sensitivity to bandwidth (BB > NB) decreased from subcortical structures to nonprimary auditory cortex, whereas sensitivity to slow modulation rates was largest in nonprimary auditory cortex and largely absent in subcortical structures. Furthermore, there was no parametric interaction between bandwidth and modulation rate. These results challenge simple hierarchical models, in that BB stimuli evoked stronger responses in primary auditory cortex (and subcortical structures) rather than nonprimary cortex. Furthermore, the strong preference for slow modulation rates in nonprimary cortex demonstrates the compelling global sensitivity of auditory cortex to modulation rates that are dominant in the principal signals that we process, e.g., speech.

hierarchical organization; cortex; sinusoidal amplitude modulation

ACOUSTIC COMMUNICATION SIGNALS, including human speech, transmit information within specific spectral bandwidths and amplitude modulation (AM) rates (Elliott and Theunissen 2009; Singh and Theunissen 2003). Both the spectral and temporal attributes of complex stimuli are thought to be processed via hierarchically organized spectral or temporal filter banks. A canonical assumption is that spectral filter width increases, and best temporal modulation rate decreases as one ascends the central auditory pathway. For example, with respect to differential spectral bandwidth sensitivity in auditory cortex, Rauschecker et al. (1995) report that the most common response profile of neurons in macaque primary auditory cortex is a decline in responsiveness with increasing stimulus bandwidth. Whereas these findings and interpretations are accurate, they may be compatible with a population response that increases with spectral bandwidth. For example, inhibitory interneurons may be more broadly tuned in auditory and other sensory cortices (Atencio and Schreiner 2008; Cardin et al. 2007; Ma et al. 2010; Simons and Carvell 1989), and their activity could provide a larger contribution to the blood-oxygen-level-dependent (BOLD) signal elicited by broad spectral bandwidth stimuli. The original model has received support from neurophysiological and imaging studies (Chevillet et al. 2011; Giraud et al. 2000; Harms and Melcher 2002, 2003; Joris et al. 2004; Rauschecker et al. 1995; Wessinger et al. 2001). If it is an accurate representation of spectral bandwidth and temporal modulation rate processing in the auditory system, then one would predict a three-way interaction of spectral bandwidth, modulation rate, and place along the auditory neuraxis. More specifically, the BOLD functional MRI (fMRI) signal should increase with increasing spectral bandwidth and decreasing modulation rate as one ascends the auditory neuraxis; early auditory areas should show reduced BOLD signal in the presence of broadband (BB) and/or slow modulation rate sounds. Here, we test this hypothesis in the human auditory system using fMRI. For a similar approach, but exploring electrophysiological response parameters, see the companion paper by Wang and colleagues (2012), using magnetoencephalography (MEG) as the recording technique.

Hierarchical models of auditory processing are often derived from processing principles in vision: as one ascends the visual hierarchy, receptive fields are sensitive to larger regions of visual space and to increasingly abstract properties of visual stimuli (Hubel and Wiesel 1968; Maunsell and Newsome 1987). By analogy, it has been proposed that the preferred spectral bandwidth of neurons or neuronal ensembles increases with ascending areas of the auditory system, in particular, at the level of cortex. For example, Rauschecker et al. (1995) tested response properties in auditory cortex of anesthetized rhesus monkeys and found that neurons in core cortex (A1, R) increased their discharge rate with decreasing spectral bandwidth, whereas neurons in the lateral belt (anterolateral, mediolateral, caudolateral) responded more vigorously to sounds with wider passbands. Building on this finding, Wessinger et al. (2001) used fMRI in humans and reported that bandpass white noise signals resulted in larger activation extents in auditory cortex than pure tones. Similar findings have been reported for rat and cat cortex (Carrasco and Lomber 2009b; Pandya et al. 2009). The intuition motivating this data pattern and perspective derives from the view that neuronal selectivity for higher-order (and perhaps combinatoric) properties of stimuli is reflected in more lowpass characteristics of cortical circuits. That is, the transformation from low-level, faithful stimulus encoding, using narrowly tuned receptive fields, to a putative encoding of complex properties over broader bandwidths and longer durations (and therefore, receptive field

Address for correspondence and present address: T. Overath, UCL Ear Institute, Univ. College London, London, WC1X 8EE, U.K. (e-mail: t.overath@ucl.ac.uk).

properties tuned to more complex signal attributes) is a principle of cortical stimulus representation.

However, bandwidth sensitivity may not be a static-response property of neuronal processing. For example, subthreshold responses (i.e., inhibitory and excitatory synaptic potentials) are much more broadly tuned than spiking responses in both midbrain (Xie et al. 2007) and cortex (Tan et al. 2004; Wehr and Zador 2003), and a neuron's suprathreshold response can vary significantly, depending on its stimulus history (Malone et al. 2002; Malone and Semple 2001). Furthermore, bandwidth sensitivity depends on the activity of both ascending and descending connections (Carrasco and Lomber 2009a, b; Pandya et al. 2009; Zhang and Suga 2000; Zhang et al. 2008; Zhou and Jen 2000), suggesting that it may be adjusted in response to the acoustic environment, the specific auditory task, and the attentional state. Each of these considerations presumably contributes to global measures of activity, such as the BOLD signal in fMRI (Goense and Logothetis 2008; Logothetis 2002; Viswanathan and Freeman 2007). For example, auditory structures in the human brainstem have been shown to display more robust responses to BB than to narrowband (NB) sounds when measured with fMRI (Hawley et al. 2005).

Similar to the concept of increasing spectral bandwidths in the auditory system, there is evidence that temporal modulation filters (Dau et al. 1997; Jepsen et al. 2008) are also hierarchically organized (Baumann et al. 2011). For example, there is a gradual reduction of synchronized firing to periodic stimuli, as the recording site moves from thalamus to primary auditory cortex to rostral field in awake marmosets (Bartlett and Wang 2007; Bendor and Wang 2007; Lu et al. 2001). Human imaging studies also reveal distinct representations of temporal modulation rates with respect to differential shapes of the hemodynamic-response function (hrf). Fast rates generally elicit strong phasic responses, whereas slower rates drive a more sustained response, particularly in auditory cortex (Giraud et al. 2000; Harms et al. 2005; Harms and Melcher 2002). To this end, Harms and Melcher (2003) developed a physiologically motivated basis set, which explicitly models different aspects of the hrf as a function of temporal modulation rate. However, the existence of an organized or hierarchical representation of modulation filters in human auditory cortex has, so far, proved elusive [see also Schönwiesner and Zatorre (2009)].

An important consideration related to the proper assessment of responses to temporal modulation concerns adaptation to long-duration stimuli. Typically, studies using sinusoidal AM (sAM) stimuli have used constant modulation rates (Joris et al. 2004). However, the response of auditory neurons often declines to an uninterrupted stimulus, as in stimulus-specific adaptation (Anderson et al. 2009; Malmierca et al. 2009; Malone and Semple 2001; Ulanovsky et al. 2004), and generally increases along the auditory system, as the neural code moves from a faithful representation of the acoustic signal to a representation of more complex stimulus attributes. Specifically, there is neural and behavioral evidence for adaptation to amplitude-modulated stimuli in the auditory system (Bartlett and Wang 2005; Wojtczak and Viemeister 2005). Taken together with the observation that naturally occurring signals, such as speech, contain a range of temporal AM rates, we examined the efficacy of constant vs. temporally jittered sAM rates in a first study so as to choose the most effective stimulus for the second, main study, comparing spectral bandwidth and temporal modulation rate in the human auditory system.

We performed two fMRI studies to characterize the response to temporally jittered vs. regular signals (*Study 1*), as well as the interaction of four different temporal modulation rates across two spectral bandwidths (*Study 2*). Our data do not support the conjecture that as bandwidth increases, higher-order areas are preferentially recruited or vice versa, that primary areas preferentially respond to spectrally narrow sounds. Instead, we report a double dissociation between subcortical and cortical structures with respect to spectral bandwidth and temporal modulation rate: subcortical structures [inferior colliculus (IC) and medial geniculate body (MGB)] are largely insensitive to the modulation rates tested here but show a high sensitivity to bandwidth, whereas cortex displays strong temporal bandpass-filter characteristics for slow modulation rates and only rather moderate sensitivity to spectral bandwidth. Crucially, the pattern of results observed for cortical structures parallels the data obtained using MEG in the companion paper by Wang and colleagues (2012).

MATERIALS AND METHODS

Participants

Four participants (one left-handed; mean age = 29; age range = 21–43; one female) took part in *Study 1*. Twenty participants took part in *Study 2*. Two participants had to be excluded because they displayed no robust activation to sounds, leaving a total of 18 participants (two left-handed; mean age = 23.83; age range = 18–34; nine females). The study was approved by the local Institutional Review Board (New York University's Committee on Activities Involving Human Subjects), and all participants provided written consent prior to participation.

Stimuli

All sounds were created digitally using Matlab 7.7 software (MathWorks, Natick, MA) at a sampling frequency of 44.1 kHz and 16-bit resolution.

Study 1: stimulus adaptation. Previous studies have generally used a constant or regular modulation rate (Harms et al. 2005; Harms and Melcher 2002; Joris et al. 2004), which may lead to stimulus-specific adaptation (Bartlett and Wang 2005; Wojtczak and Viemeister 2005). Therefore, the first study examined the relative efficacy of constant vs. temporally jittered sAM rates. Constant sAM sounds had a modulation rate of 3 Hz, whereas for the jittered sAM stimuli, the mean 3-Hz modulation rate within a sound was varied, such that the modulation periods (in ms) of the sinusoidal modulations were normally distributed around the mean modulation rate with a SD of one-sixth of the mean modulation period (Fig. 1). The depth of the modulation envelope was 80%. The modulated carrier was either a sinusoid (Sin), a NB, or a BB sound. The NB and BB carriers were pink noise so as to approximate equal intensity/octave. BB sounds spanned 4 octaves and were centered on 1,200 Hz (range: 300–4,800 Hz); NB sounds had a range of 0.25 octave and were centered on 600, 1,200, or 2,000 Hz (ranges: 550–654, 1,100–1,309, 1,834–2,181 Hz, respectively) to approximate the overall range of the BB sounds. The bandwidth parameters were chosen so as to represent the smallest spectral region possible (Sin), remain within one critical band (NB), or span several critical bands (BB). Thirty exemplars of each of the six stimulus conditions (three carrier types and two sAM modulation types) were created. Sounds were 6 s long, with 200 ms-raised cosine on and off ramps. Each sound was unique, both with respect to the precise pink-noise frequency spectrum of NB and BB stimuli, as well as the composition of the actual modulation periods.

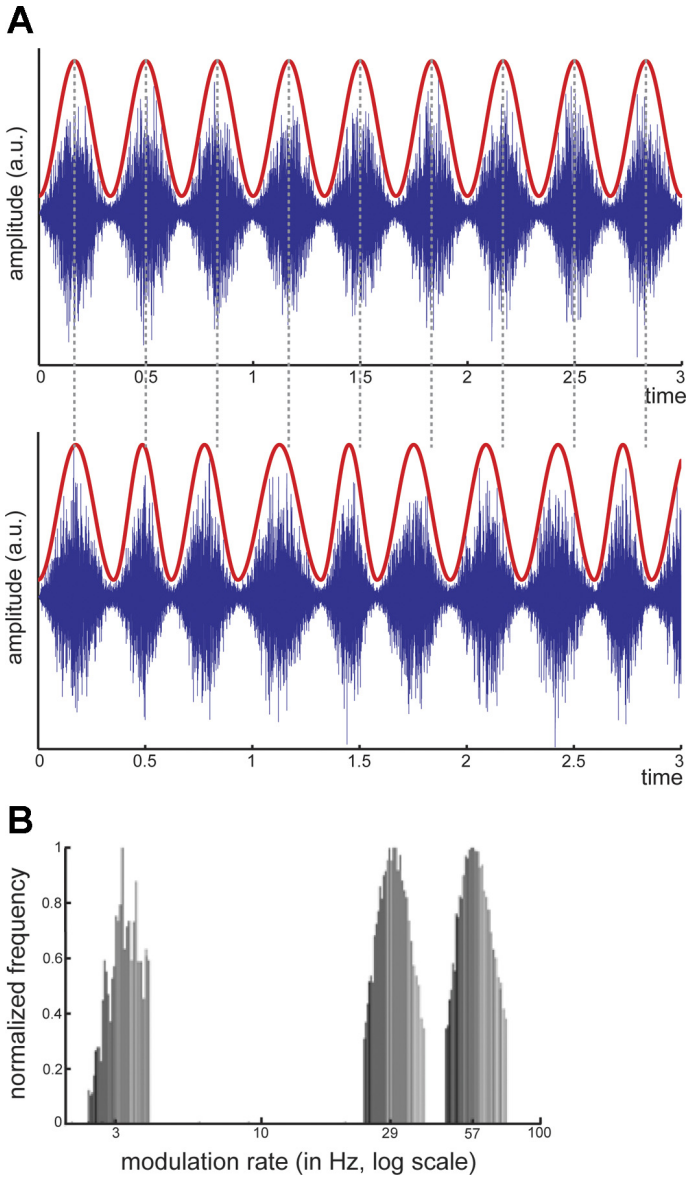


Fig. 1. *A*, *Top*: 3-s excerpt of a broadband (BB) stimulus with regular 3-Hz sinusoidal amplitude modulation (sAM). *Bottom*: 3-s excerpt of a BB stimulus with jittered mean modulation rate of 3 Hz sAM. Note the temporally jittered (variable) 1/2-period lengths of the individual modulations (cf. the dashed, vertical lines). Regular and jittered sAM stimuli were used in *Study 1*; in *Study 2*, only temporally jittered stimuli were used. *B*: normalized bar histogram distributions of the temporally jittered sAM stimuli with mean rates of 3, 29, and 57 Hz.

Study 2: spectral and temporal sensitivity. In the second, main study, the stimuli were the jittered NB and BB sounds, as above, this time with the use of four (0, 3, 29, 57 Hz) mean sAM rates (see Fig. 1). Forty exemplars for each of the eight conditions (two bandwidth types and four AM types) were created (for NB sounds; 13, 14, and 13 exemplars were centered on 600, 1,200, and 2,000 Hz, respectively). We chose NB sounds instead of Sin (as in *Study 1*) in an effort to render the stimuli more ecologically valid, as bandpassed signals are more common in nature than isolated sinusoids. Each sound was unique, both with respect to the precise pink-noise frequency spectrum of NB and BB stimuli, as well as the composition of the actual modulation periods.

The bandwidth parameters were chosen so as to stay within one critical band (NB) or span several critical bands (BB). The modulation rate parameters were chosen so as to approximate characteristic AM rates in natural signals, such as speech, e.g., delta/theta and gamma for 3 Hz and 29 Hz stimuli, respectively (Pickett 1999; Rosen 1992), while at the same time including a rate just above the pitch threshold (57 Hz) (Pressnitzer et al. 2001).

For the comparison of NB and BB stimuli, one concern was whether to balance overall signal intensity (root-mean-square) or perceived loudness across stimuli. One step to address this was the choice of pink noise, which has equal intensity/octave and should therefore decrease the effect of increasing loudness with increasing bandwidth. Pilot psychophysical measures (not shown) showed that participants were often unable to provide a reliable assessment of relative loudness when comparing signals differing in more than one dimension, i.e., modulation rate and spectral bandwidth (e.g., 3 Hz AM of BB sound vs. 57 Hz AM of NB sound). Taken together with the observation that equal energy and equal spectrum level sounds do not elicit significantly different responses in subcortical structures (Hawley et al. 2005) and that previous studies show an effect of signal energy only for much larger differences than present in our study (Langers et al. 2007), we decided to present stimuli at equal signal intensity.

Experimental design

Study 1: stimulus adaptation to modulation type. We used a 3 bandwidth (Sin, NB, BB) × 2 modulation type (constant, jittered) design. Stimuli were presented randomly over three sessions. Each session consisted of 72 scans, where each experimental condition and a silent baseline condition were presented 10 times (two dummy scans were acquired at the beginning of each session to control for image saturation effects). To ensure and control that participants were alert, they were asked to keep their eyes open and to press a button at the beginning of each scan. The stimuli were presented at a comfortable listening level (~65–70 dB sensation level) via Sensimetrics (Malden, MA) MRI-compatible insert earphones (Model S14); in addition, participants wore protective earmuffs to further reduce the background noise of the scanner environment.

Study 2: spectral and temporal sensitivity. We used a 2 bandwidth (BB, NB) × 4 modulation rate (0, 3, 29, 57 Hz) design. The stimuli were presented randomly over four sessions. Each session consisted of 92 scans, in which each experimental condition and a silent baseline condition were presented 10 times (plus two dummy scans at the beginning of the session). The task and stimulus presentation were the same as in the first study.

Image acquisition

Imaging parameters were the same for the two studies. T2* gradient-weighted echo-planar images (EPIs) were acquired on a 3 Tesla Siemens Allegra system (Erlangen, Germany) using a Nova Medical NM-011 head coil (Wilmington, MA). Thirty slices (2 × 2 × 2 mm voxels) were acquired with each scan, using a sparse acquisition paradigm (Hall et al. 1999), where each volume was cardiac gated [minimum time to repeat (TR)/time to echo (TE): 8,910/30 ms]; this ensured improved imaging of subcortical auditory structures (IC and MGB), which are particularly susceptible to cardiac pulse motion artifacts. The acquisition volume was tilted forward, such that slices were parallel to and centered on the superior temporal gyrus. A structural high-resolution, T1-weighted MRI magnetization-prepared rapid gradient echo scan was acquired for each participant at the end of the experimental session.

Data analysis

The principles of data analysis were the same in the two studies. Imaging data were analyzed using Statistical Parametric Mapping software (SPM5, http://www.fil.ion.ucl.ac.uk/spm). The first two volumes in each session were discarded to control for T1 saturation

effects. The remaining scans were realigned to the first volume in the first session, unwarped to correct for motion artifacts, and resliced using sinc interpolation (SPM5, "Realign & Unwarp"); the structural scan of each participant was coregistered to the mean functional scan (SPM5, "Coregister") and segmented into gray and white matter and cerebrospinal fluid and spatially normalized (SPM5, "Segment") before applying the resulting transformations to the EPIs and structural scan (SPM5, "Normalize: Write"). Finally, the EPIs were spatially smoothed to improve the signal-to-noise ratio: for the analysis of cortical structures, data were smoothed using an isotropic, 8-mm full-width at half-maximum (FWHM) Gaussian kernel; for the analysis of subcortical structures (IC, MBG), data were smoothed with an isotropic, 4-mm FWHM Gaussian kernel.

Since the data were acquired using a cardiac-gated procedure, the actual TR and the precise time of acquisition after the end of each sound varied slightly for each scan. To model the hemodynamic response to the stimuli most accurately with respect to this variation, the duration of each sound was specified in relation to the length of its respective TR. For example, in *Study 2*, the mean TR across participants was 9,380 ms, with the range of the mean TRs spanning 9,270–9,500 ms. The stimulus box-car functions were then convolved with the canonical hrf.

The statistical analyses of *Study 1* and *Study 2* are based on fixed-effects and random-effects models, respectively, within the context of the general linear model (Friston et al. 1995), using a statistical threshold of $P < 0.05$, corrected for family-wise error (FWE; SPM5, one-sample *t*-test, and SPM8, one-way ANOVA, within subject). Regions of interest (ROIs) in cortex were anatomically informed and were based on 30% probability maps of the three sub-areas of Heschl's gyrus (HG; Te1.0, Te1.1, Te1.2) (Morosan et al. 2001) within the SPM Anatomy Toolbox (Eickhoff et al. 2005) and planum temporale (PT) (Westbury et al. 1999). ROIs of IC and MGB were defined functionally by forming spheres with a 4-mm radius around the [*x y z*]coordinates of activation maxima in these structures. In the first study, these were for left and right IC [−4 −34 −10] and [4 −36 −8] and for left and right MGB [−16 −26 −6] and [16 −26 −6], respectively. In the second study, for IC, the spheres were centered on [−4 −36 −8] and [6 −36 −8], whereas for MGB they were centered on [−16 −26 −4] and [18 −26 −6] for left and right hemispheres, respectively. These correspond closely to previously reported coordinates of these structures (Devlin et al. 2006). ROI and activation-extent analyses were performed using MarsBaR (http://marsbar.sourceforge.net/) (Brett et al. 2002) and SPSS (IBM, Armonk, NY).

RESULTS

We report three findings. First, temporal modulation type (jittered vs. regular) shows a compelling effect of jittering on the BOLD signal across spectral bandwidths. Second, the effect of bandwidth yields effects contrary to the initial assumption, not consistent with a processing hierarchy-bandwidth coupling. Third, the data demonstrate a striking sensitivity to slow temporal modulation rates in auditory cortex.

Two metrics to analyze the BOLD fMRI data are used: activation strength (i.e., the statistic at a given voxel) and activation extent (i.e., the number of contiguous voxels that are above a chosen statistical threshold). These two metrics reflect different preferences: activation strength judges the significance of an effect at each individual voxel and therefore is largely independent of activation extent; in contrast, activation extent judges the significance of an effect on its spatial extent and is independent of the activation strength. Activation strength is more widely used, whereas some statistics combine the two to provide an integrated and more lenient, statistical metric (e.g., cluster-based statistics). Both are used here, since previous studies directly relevant to the present findings have reported activation-extent metrics [e.g., Chevillet et al. (2011) and Wessinger et al. (2001)].

Effect of varying modulation rate type at different carrier bandwidths (Study 1)

Activation strength. To determine whether there was response adaptation, the first study compared amplitude-modulated stimuli that were constant in modulation rate with those in which rate varied around a mean modulation rate. The cardiac-gated image acquisition allowed us to investigate these effects in both subcortical and cortical areas. The response strength was larger for jittered compared with constant modulation rates. We computed a four-way repeated-measures ANOVA with factors ROI (IC, MGB, Te1.0, Te1.1, Te1.2, PT), hemisphere (left, right), modulation type (jittered, regular), and bandwidth (Sin, NB, BB). The overall greater activation strengths in cortex were reflected in a main effect of ROI [$F_{(5,15)} = 15.1$, $\eta^2_p = 0.83$], while a main effect of hemisphere [$F_{(1,3)} = 16.28$, $\eta^2_p = 0.84$] revealed stronger activations in the right hemisphere. The jittered sAM sounds yielded stronger effects [main effect of modulation type, $F_{(1,3)} = 18.78$, $\eta^2_p = 0.86$], while activity increased as a function of bandwidth [$F_{(1.01,3.02)} = 50.31$, $\eta^2_p = 0.94$; in cases where sphericity cannot be assumed, degrees of freedom are adjusted using the Greenhouse-Geisser correction].

There were ROI × modulation type [$F_{(5,15)} = 6.32$, $\eta^2_p = 0.68$] and ROI × bandwidth [$F_{(10,30)} = 7.5$, $\eta^2_p = 0.71$] interactions, as well as a trend for a modulation type × bandwidth interaction [$F_{(2,6)} = 3.82$, $P = 0.09$, $\eta^2_p = 0.56$]. None of the other interactions reached significance; importantly, there was no ROI × modulation type × bandwidth interaction.

To investigate the effects of modulation type and bandwidth in more detail in each ROI, we computed three-way repeated-measures ANOVAs with factors hemisphere, modulation type, and bandwidth. For all ROIs, there were main effects of bandwidth and modulation type (except for IC; see Table 1 and Fig. 2). The ANOVAs revealed no main effects of hemisphere and no interactions between any of the factor combinations (e.g., no modulation type × bandwidth interaction).

Activation extent. The analysis of modulation type variation and bandwidth was also evaluated with respect to the activation-extent metric. To this end, we contrasted each condition with the silent baseline condition and determined the number of voxels that displayed significant activation in auditory cortex (voxels had to be within the superior temporal lobule, extending into the superior temporal sulcus) at a threshold of $P < 0.05$ [FWE corrected across the whole volume; e.g., see Wessinger et al. (2001)].

We performed two types of activation-extent analyses. First, we determined the overall activation extent in auditory cortex. With respect to bandwidth, one would expect larger activation extents as spectral bandwidth increases; more specifically, as one ascends the neuraxis, spectral bandwidth sensitivity should change: neurons in core cortical regions should respond, not at all or weakly to BB sounds but strongly to pure tones (Rauschecker et al. 1995), whereas nonprimary cortex should respond increasingly to BB sounds.

Table 1. *Study 1*

Activation Strength

	Bandwidth (Sin, NB, BB)		Mod Type (jitter, constant)	
	$F(2,6)$	η^2_p	$F(1,3)$	η^2_p
IC	15.16*	0.86	n.s.	–
MGB	10.42	0.78	10.90	0.78
Te1.0	81.40	0.96	18.81	0.86
Te1.1	52.23	0.95	12.60	0.80
Te1.2	32.02	0.91	14.39	0.83
PT	11.21†	0.79	46.81	0.94

Percentage Activation Extent

	Bandwidth (Sin, NB, BB)		Mod Type (jitter, constant)	
	$F(2,6)$	η^2_p	$F(1,3)$	η^2_p
Te1.0	49.00	0.94	n.s.	–
Te1.1	35.55	0.92	*9.59*	*0.76*
Te1.2	17.99	0.86	n.s.	–
PT	*6.67*‡	*0.69*	*6.06*	*0.67*

Activation Strength: main effects of bandwidth (columns 2 and 3) and modulation (Mod) type (columns 4 and 5) with respect to F-statistic and partial η^2_p in the regions of interest (ROIs; column 1). The effect size η^2_p (Cohen 1973; Pierce et al. 2004) describes the ratio of variance accounted for by the treatment effect divided by variance accounted for by the treatment effect plus its associated error variance; to a 1st approximation, $\eta^2_p \cdot 100$ denotes the percentage variance that is attributable to the factor. Adjusted degrees of freedom, according to the Greenhouse-Geisser correction, are displayed where appropriate: *F(1.02,3.05); †F(1.01,3.03). Sin, sinusoid; NB, narrowband; BB, broadband; IC, inferior colliculus; MGB, medial geniculate body; PT, planum temporale; n.s., not significant. Percentage Activation Extent: main effects of bandwidth (columns 2 and 3) and modulation type (columns 4 and 5) with respect to F-statistic and partial η^2_p in the cortical ROIs (column 1). Adjusted degrees of freedom, according to the Greenhouse-Geisser correction, are displayed where appropriate: ‡F(1.02,3.06). Italics denote trend-level significance ($P < 0.1$).

Second, we determined the percentage activation extent within the cortical ROIs. Whereas the first analysis provides an overall measure of activation extent without detailed functional and anatomical constraints, the second analysis provides a more focal measure of activation extent within anatomically defined ROIs in auditory cortex. For example, for the effect of spectral bandwidth, one would expect the percentage activation extent to decrease for NB sounds from primary to nonprimary cortex, whereas BB sounds should show the reverse trend.

With respect to the overall activation extent in auditory cortex, a three-way repeated-measures ANOVA with factors hemisphere, modulation type, and bandwidth revealed main effects of hemisphere [$F_{(1,3)} = 11.83$, $\eta^2_p = 0.8$], modulation type [$F_{(1,3)} = 13.88$, $\eta^2_p = 0.82$], and bandwidth [$F_{(2,6)} = 14.22$, $\eta^2_p = 0.8$]. That is, the activation extent was generally, significantly larger in the right hemisphere and for jittered sAM sounds. Furthermore, activation extent increased with spectral bandwidth. The results also showed a hemisphere × modulation type interaction [$F_{(1,3)} = 12.3$, $\eta^2_p = 0.8$]; pairwise comparisons revealed that Sin carriers produced significantly larger activation extents in right auditory cortex ($P < 0.05$, Bonferroni corrected), whereas there were similar trends for NB and BB sounds ($P = 0.05$, and $P = 0.08$, respectively).

Figure 3 visualizes the overall activation extent for Sin, NB, and BB carriers across modulation types in cortex with respect to areas that displayed significantly stronger activation for increasing bandwidth (i.e., Sin < NB < BB; see activation-strength results above). At the group level, the activation patterns of the three sAM carriers were overlapping in much of primary cortex (Fig. 3). However, much of HG showed significantly stronger responses as a function of bandwidth.

For the second analysis, we first performed an overall four-way repeated-measures ANOVA with factors ROI (Te1.0, Te1.1, Te1.2, PT), hemisphere, modulation type, and bandwidth, which revealed a main effect of bandwidth [$F_{(2,6)} = 26.48$, $\eta^2 = 0.9$] and trends for ROI [$F_{(3,9)} = 2.92$, $P = 0.09$, $\eta^2 = 0.49$], hemisphere [$F_{(1,3)} = 8.73$, $P = 0.06$, $\eta^2 = 0.74$], and modulation type [$F_{(1,3)} = 6.83$, $P = 0.08$, $\eta^2 = 0.7$]. The only significant interaction was a ROI × bandwidth interaction [$F_{(6,18)} = 9.45$, $\eta^2 = 0.76$]; pairwise comparisons revealed significant differences ($P < 0.05$, Bonferroni corrected) between BB and both Sin and NB sounds in Te1.0 and Te1.1, but no significant differences for these comparisons in Te1.2 or PT.

We also investigated the effects in more detail for each cortical ROI using a three-way repeated-measures ANOVA with factors hemisphere, modulation type, and bandwidth. There were significant main effects of bandwidth in all ROIs (Sin showed a trend toward significance), as well as trends for modulation type in Te1.1 and PT, but no main effects of hemisphere and no interactions (see Table 1 and Fig. 3).

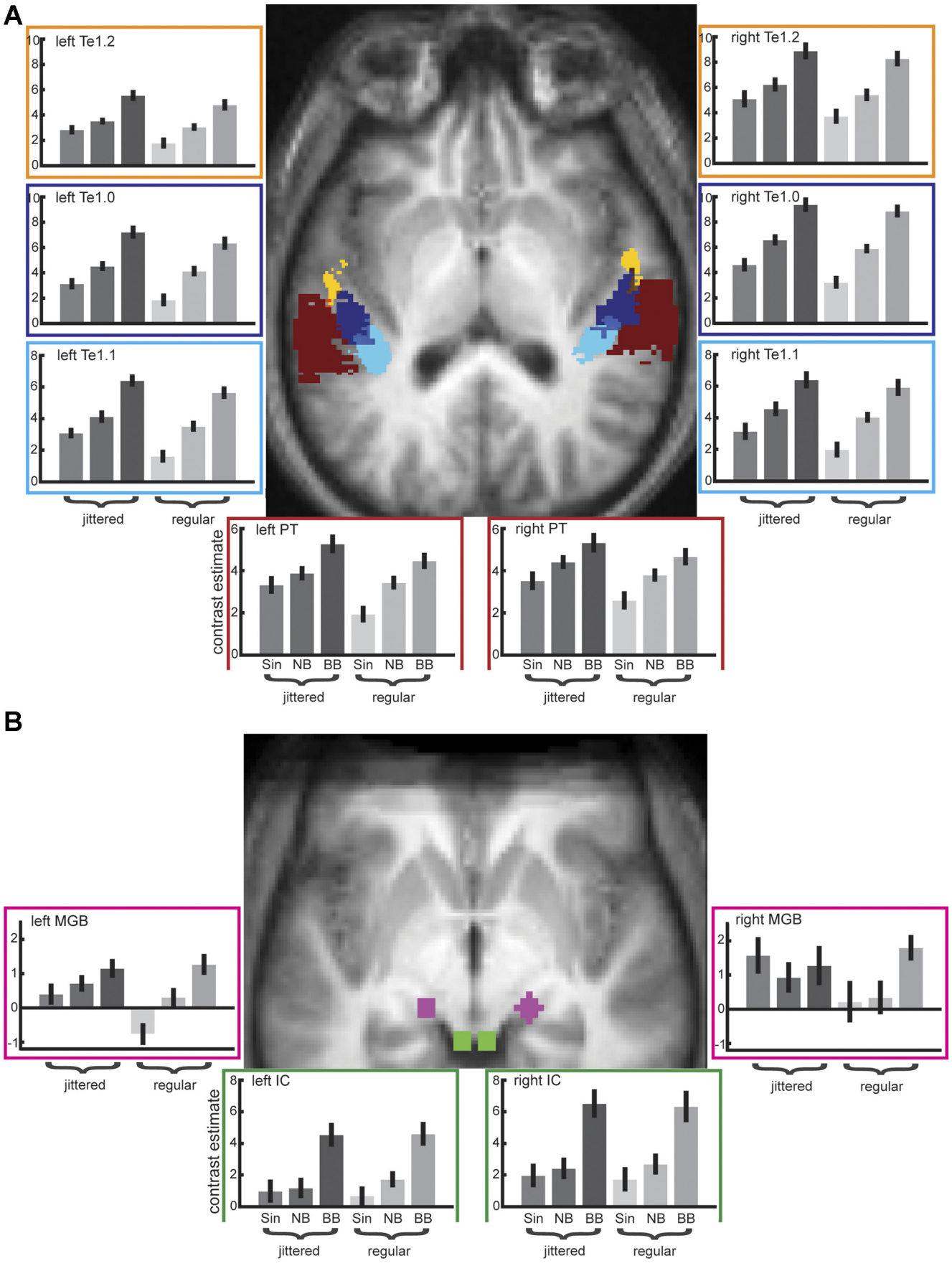
Even at the current, comparatively high voxel resolution of 2 × 2 × 2 mm, IC and MGB only encompass a few voxels [for IC, e.g., see Hawley et al. (2005)]; volumetric analyses of fMRI activation extents in these structures are therefore not particularly meaningful and are not pursued here.

In conclusion, since the jittered sAM stimulus proved to be more efficacious for both activation-strength and activation-extent analyses, we used this stimulus in the subsequent study of bandwidth and modulation rate (*Study 2*).

Spectral bandwidth and temporal modulation rate response characteristics (Study 2)

Activation strength. Results are presented in the same manner and order as in *Study 1*. The contrast estimates for the eight stimulus conditions yielded distinct response patterns in the six ROIs (Fig. 4). A four-way repeated-measures ANOVA with factors ROI (IC, MGB, Te1.0, Te1.1, Te1.2, PT), hemisphere (left, right), bandwidth (NB, BB), and modulation rate (0, 3, 29, 57) revealed main effects for ROI [$F_{(2.41,40.95)} = 22.87$, $\eta^2_p = 0.57$], hemisphere [$F_{(1,17)} = 10.43$, $\eta^2_p = 0.38$], bandwidth [$F_{(1,17)} = 150.31$, $\eta^2_p = 0.90$], and modulation rate [$F_{(3,51)} = 37.28$, $\eta^2_p = 0.69$]. Furthermore, there were ROI × bandwidth [$F_{(2.65,44.99)} = 73.8$, $\eta^2_p = 0.83$] and ROI ×

Fig. 2. Activation-strength analysis in *Study 1*. *A*: cortical regions of interest (ROIs). The differently shaded gray bars plot contrast estimates [±90% confidence interval (CI)] of the 6 experimental conditions in the 30% probability map ROIs in bilateral Te1.0 (purple), Te1.1 (cyan), Te1.2 (orange), and planum temporale (PT; red). Sin, sinusoid; NB, narrowband; BB, broadband. *B*: subcortical ROIs. The differently shaded gray bars plot contrast estimates (±90% CI) of the 6 experimental conditions in the functionally defined ROIs in bilateral inferior colliculus (IC; green) and medial geniculate body (MGB; magenta).


A
left Te1.2
left Te1.0
left Te1.1
right Te1.2
right Te1.0
right Te1.1
jittered
regular
left PT
right PT
contrast estimate
Sin NB BB Sin NB BB
jittered
regular
B
left MGB
right MGB
left IC
right IC
contrast estimate
jittered
regular

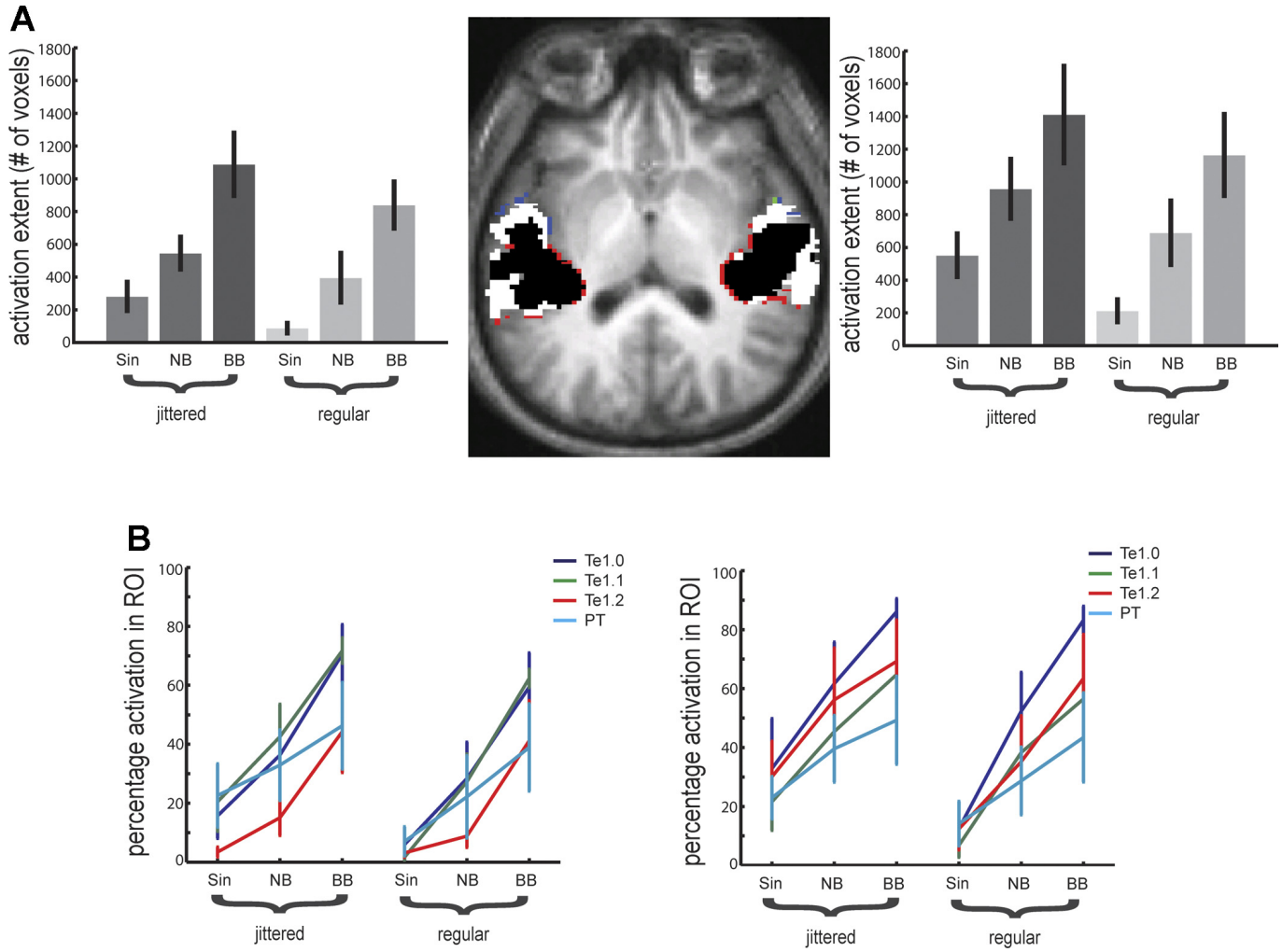


Fig. 3. Activation-extent analyses in *Study 1*. *A*: the bar plots display the overall mean activation extent (±SE) for the 6 experimental conditions in left and right auditory cortex, respectively. The overlay image in the middle depicts a visual combination of overall activation-extent and activation-strength analyses. Color code: white (Sin and NB and BB); green (Sin only); blue (NB only); red (BB only); black (bandwidth increase: Sin < NB < BB). Note that in areas coded in green (Sin only), blue (NB only), and red (BB only), the activation strengths among Sin, NB, and BB are not statistically different from each other; only the areas in black reveal a statistically stronger response with increasing spectral bandwidth. Results are superimposed on a tilted (pitch: −0.5) axial section of the 4 participants' average structural scans. *B*: mean percentage activation extent (±SE) in the cortical ROIs Te1.0 (purple), Te1.1 (green), Te1.2 (red), and PT (cyan) in the 6 experimental conditions.

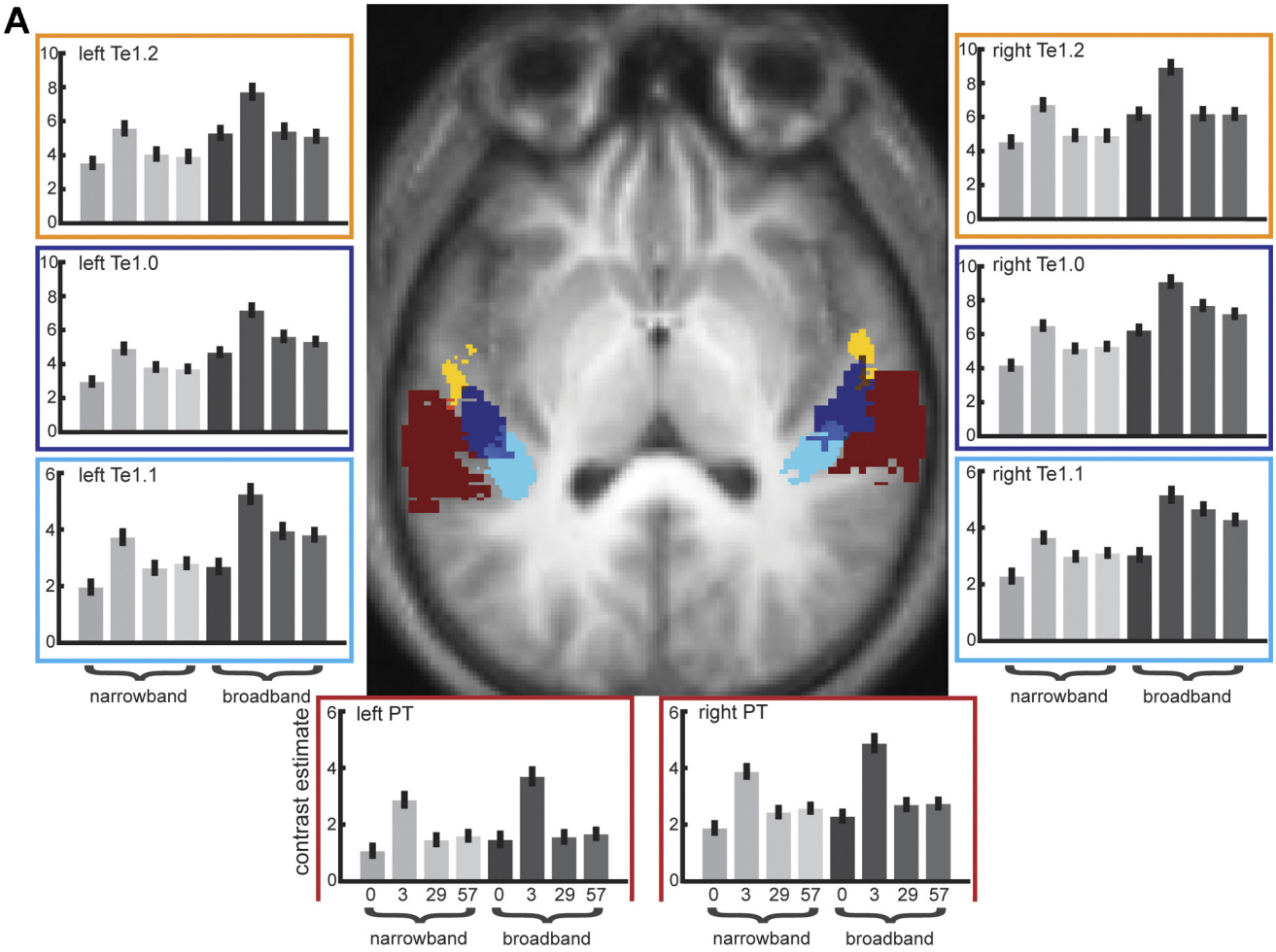
modulation rate [$F_{(4.6,78.21)} = 20.39$, $\eta^2_p = 0.55$] interactions, reflecting the overall decrease and increase of the factors bandwidth and modulation rate, respectively, across ROIs (see also below). There was a trend-level ROI × bandwidth × modulation rate interaction when using the Greenhouse-Geisser correction for the violation of sphericity [$F_{(4.80,81.61)} = 2.07$, $P = 0.08$, $\eta^2_p = 0.11$]; when using the less-conservative Huynh-Feldt correction, the interaction was only marginally not significant [$F_{(6.92,11.62)} = 2.07$, $P = 0.05$, $\eta^2_p = 0.11$].

We further investigated the response to bandwidth and modulation rate in each ROI individually by computing three-way repeated-measures ANOVAs with factors hemisphere, bandwidth, and modulation rate. All ROIs showed a main effect of bandwidth and modulation rate (with the exception of IC; see Table 2 and Fig. 4). Overall, the effect sizes for bandwidth and modulation rate decreased or increased, respectively, from subcortical structures to nonprimary auditory cortex. Whereas the response across modulation rates was relatively flat in the IC, there was more variability in the MGB. The overall BOLD signal strength was also much lower in MGB, confirming the general difficulty to obtain a robust BOLD signal in this center of the auditory pathway [e.g., Baumann et al. (2010); see also the MGB results for *Study 1*].

Activations in MGB, Te1.0, and PT were generally larger in the right hemisphere, and this was reflected by a main effect of hemisphere in these structures [$F_{(1,17)} = 5.49$, $\eta^2_p = 0.24$; $F_{(1,17)} = 19.1$, $\eta^2_p = 0.53$; $F_{(1,17)} = 18.93$, $\eta^2_p = 0.53$, respectively].

All cortical ROIs, with the exception of Te 1.0, showed bandwidth × modulation rate interactions, describing the ROI × bandwidth × modulation rate interaction for the four-

Fig. 4. Activation-strength analysis in *Study 2*. *A*: cortical ROIs. The differently shaded gray bars plot contrast estimates (±90% CI) of the 8 experimental conditions in the 30% probability map ROIs in bilateral Te1.0 (purple), Te1.1 (cyan), Te1.2 (orange), and PT (red). *B*: subcortical ROIs. The differently shaded gray bars plot contrast estimates (±90% CI) of the 8 experimental conditions in the functionally defined ROIs in bilateral IC (green) and MGB (magenta).


A
left Te1.2
left Te1.0
left Te1.1
right Te1.2
right Te1.0
right Te1.1
narrowband
broadband
left PT
right PT
contrast estimate
0 3 29 57 0 3 29 57
narrowband
broadband

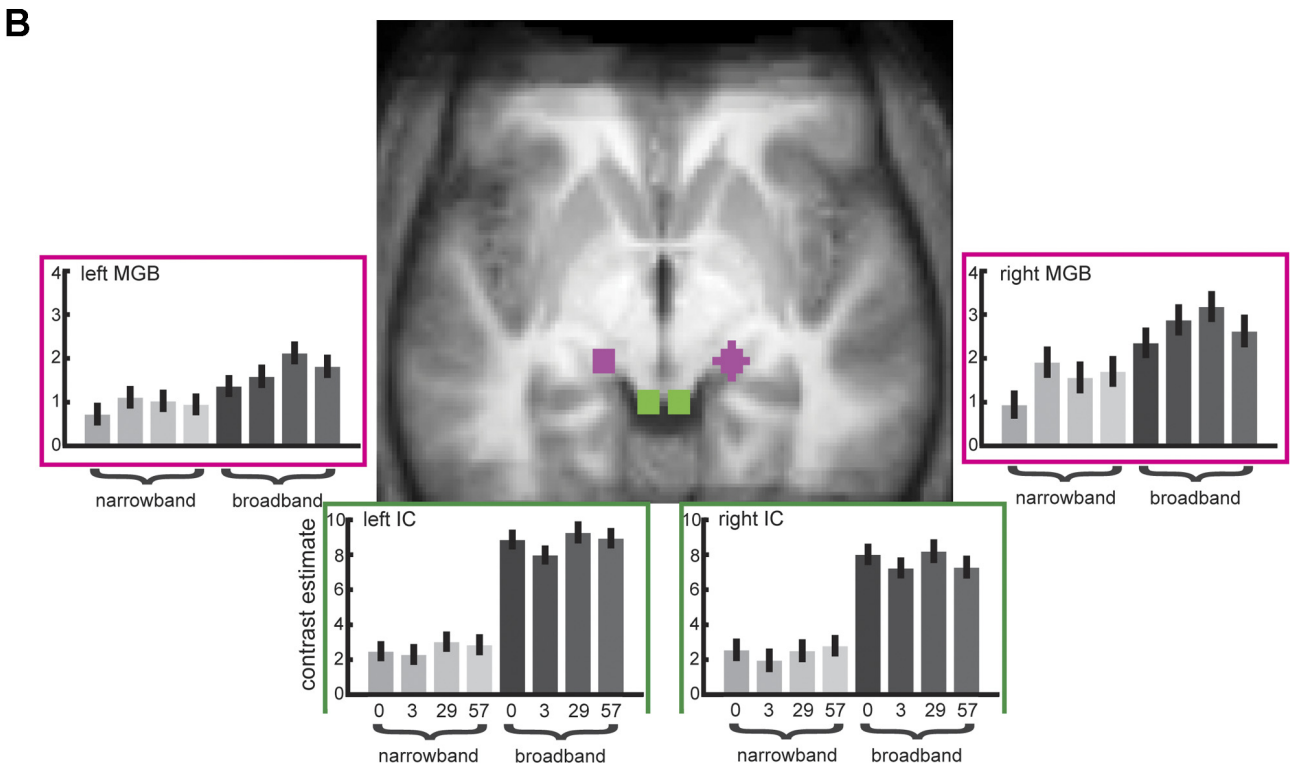

B
left MGB
right MGB
narrowband
broadband
left IC
right IC
contrast estimate
0 3 29 57 0 3 29 57
narrowband
broadband

Table 2. *Study 2*

Activation Strength

	Bandwidth (NB, BB)		Mod Rate (0, 3, 29, 57)		Bandwidth × Mod Rate	
	$F(1,17)$	η^2_p	$F(3,51)$	η^2_p	$F(3,51)$	η^2_p
IC	317.06	0.95	n.s.	–	n.s.	–
MGB	99.01	0.85	6.28	0.27	n.s.	–
Te1.0	55.80	0.77	50.02	0.75	n.s.	–
Te1.1	37.39	0.69	47.70	0.74	4.09	0.19
Te1.2	17.59	0.51	52.55	0.76	3.53	0.17
PT	6.61	0.27	72.24*	0.81	3.28†	0.16

Percentage Activation Extent

	Bandwidth (NB, BB)		Mod Rate (0, 3, 29, 57)		Bandwidth × Mod Rate	
	$F(1,17)$	η^2_p	$F(3,51)$	η^2_p	$F(3,51)$	η^2_p
Te1.0	22.61	0.61	45.19	0.73	n.s.	–
Te1.1	37.07	0.69	43.12	0.72	6.17‡	0.27
Te1.2	17.44	0.51	62.32	0.79	*2.69*	*0.14*
PT	20.27	0.54	41.48§	0.71	7.11¶	0.30

Activation Strength: main effects of bandwidth (columns 2 and 3) and modulation rate (columns 4 and 5), as well as their interaction (columns 6 and 7) with respect to F-statistic and partial η^2_p in the ROIs (column 1). Adjusted degrees of freedom, according to the Greenhouse-Geisser correction, are displayed where appropriate: *F(2.34,34.03); †F(2.97,47.43). Percentage Activation Extent: main effects of bandwidth (columns 2 and 3) and modulation rate (columns 4 and 5), as well as their interaction (columns 6 and 7) with respect to F-statistic and partial η^2_p in the cortical ROIs (column 1). Adjusted degrees of freedom, according to the Greenhouse-Geisser correction, are displayed where appropriate: ‡F(1.90,32.22); §F(1.21,20.61); ¶F(1.85,31.38). Italics denote trend-level significance ($P < 0.1$).

way ANOVA above (Table 2). Post hoc pairwise comparisons revealed significant differences between bandwidths for all modulation rates (all $P < 0.05$, Bonferroni corrected) in all ROIs except PT, which only showed a bandwidth difference for 3 Hz ($P < 0.05$). With respect to modulation rate, pairwise comparisons revealed some statistically significant differences between modulation rates within a bandwidth, but the pattern did not show a clear trend and was not consistent between ROIs. That is, contrary to the parametric effect of spectral bandwidth (see also *Study 1*), the modulation rates used here revealed a categorical rather than parametric effect, with 3 Hz sAM evoking the largest response.

Activation extent. We computed a three-way repeated-measures ANOVA with factors hemisphere, bandwidth, and modulation rate to determine an overall effect of activation extent in cortex. The results revealed main effects for all three factors, respectively [$F_{(1,17)}$, 22.67, $\eta^2_p = 0.57$; $F_{(1,17)} = 50.39$, $\eta^2_p = 0.75$; $F_{(1.33,22.67)} = 62.34$, $\eta^2_p = 0.79$]. There was also a bandwidth × modulation rate interaction [$F_{(1.94,32.92)} = 8.60$, $\eta^2_p = 0.34$]; pairwise comparisons revealed significant differences between 0 Hz and both 29 and 57 Hz for BB sounds ($P < 0.05$, Bonferroni corrected), whereas these modulation rates were not statistically different for NB sounds. All pairwise comparisons with 3 Hz were significant ($P < 0.05$, Bonferroni corrected).

This analysis allowed us to visualize the overall activation pattern for NB and BB sounds in cortex with respect to areas that showed significantly stronger activations for BB than for NB sounds (BB > NB). Figure 5 shows that at the group level, BB and NB sounds led to largely overlapping activations in much of primary cortex (Fig. 5). Furthermore, only primary cortex showed significantly stronger responses to BB than to NB sounds, reflecting the fact that Te1.0 had the largest effect size for bandwidth in cortex.

With respect to the percentage of activated voxels within the cortical ROIs, we performed a four-way repeated-measures ANOVA with factors ROI (Te1.0, Te1.1, Te1.2, PT), hemisphere, bandwidth, and modulation rate. The results revealed main effects for ROI [$F_{(2.02,34.38)} = 19.04$, $\eta^2_p = 0.53$], hemisphere [$F_{(1,17)} = 15.36$, $\eta^2_p = 0.48$], bandwidth [$F_{(1,17)} = 52.00$, $\eta^2_p = 0.75$], and modulation rate [$F_{(2.50,42.45)} = 77.38$, $\eta^2_p = 0.82$]. Furthermore, there were ROI × bandwidth [$F_{(1.64,27.92)} = 10.44$, $\eta^2_p = 0.38$], ROI × modulation rate [$F_{(4.21,71.51)} = 11.20$, $\eta^2_p = 0.40$], and bandwidth × modulation rate [$F_{(3,51)} = 3.40$, $\eta^2_p = 0.17$] interactions, as well as a ROI × bandwidth × modulation rate [$F_{(4.59,78.01)} = 3.69$, $\eta^2_p = 0.18$] interaction.

We also investigated each cortical ROI in more detail via separate three-way repeated-measures ANOVAs with factors hemisphere, bandwidth, and modulation rate. These revealed main effects of bandwidth and modulation rate in all ROIs and bandwidth × modulation rate interactions in all ROIs except Te1.0 (trend-level significance in Te1.2; Table 2, Percentage Activation Extent). Pairwise comparisons revealed that in all ROIs, BB sounds led to larger activation extents for each temporal modulation rate (all $P < 0.05$, Bonferroni corrected).

No area displayed a significantly greater response (or a greater activation extent) to NB—compared with BB—sounds for any modulation rate, even when the threshold was lowered to $P < 0.05$ (uncorrected for multiple comparisons).

DISCUSSION

The fMRI data that we present speak to three basic issues. First, we asked whether jittering the rate of AM yields more robust activation. Our data suggest that jittering enhances response magnitude, possibly by mitigating against stimulus-specific adaptation. Second, we investigated the extent to which bandwidth sensitivity increases as one ascends the auditory pathway, a standard assumption in hierarchical processing models. We did not find such a pattern. The present findings revealed stronger responses to BB compared with NB, sounds throughout the human central auditory system. However, this effect was strongest in subcortical structures (IC and MGB) and decreased from primary to nonprimary auditory cortices. Third, we tested different temporal modulation rates, known to lie in the regions that play a dominant role in perceptual analysis of auditory communication signals. We report a heightened sensitivity to low-rate modulation, consistent with the lowpass characteristics of cortex, a finding that is reinforced in the companion MEG paper (Wang et al. 2012). In particular, nonprimary auditory cortex (i.e., PT) displayed the strongest response to the slow modulation rate tested (3 Hz), whereas this effect decreased in primary auditory areas and was absent in IC. Critically, we did not observe a parametric interaction between bandwidth and modulation rate.

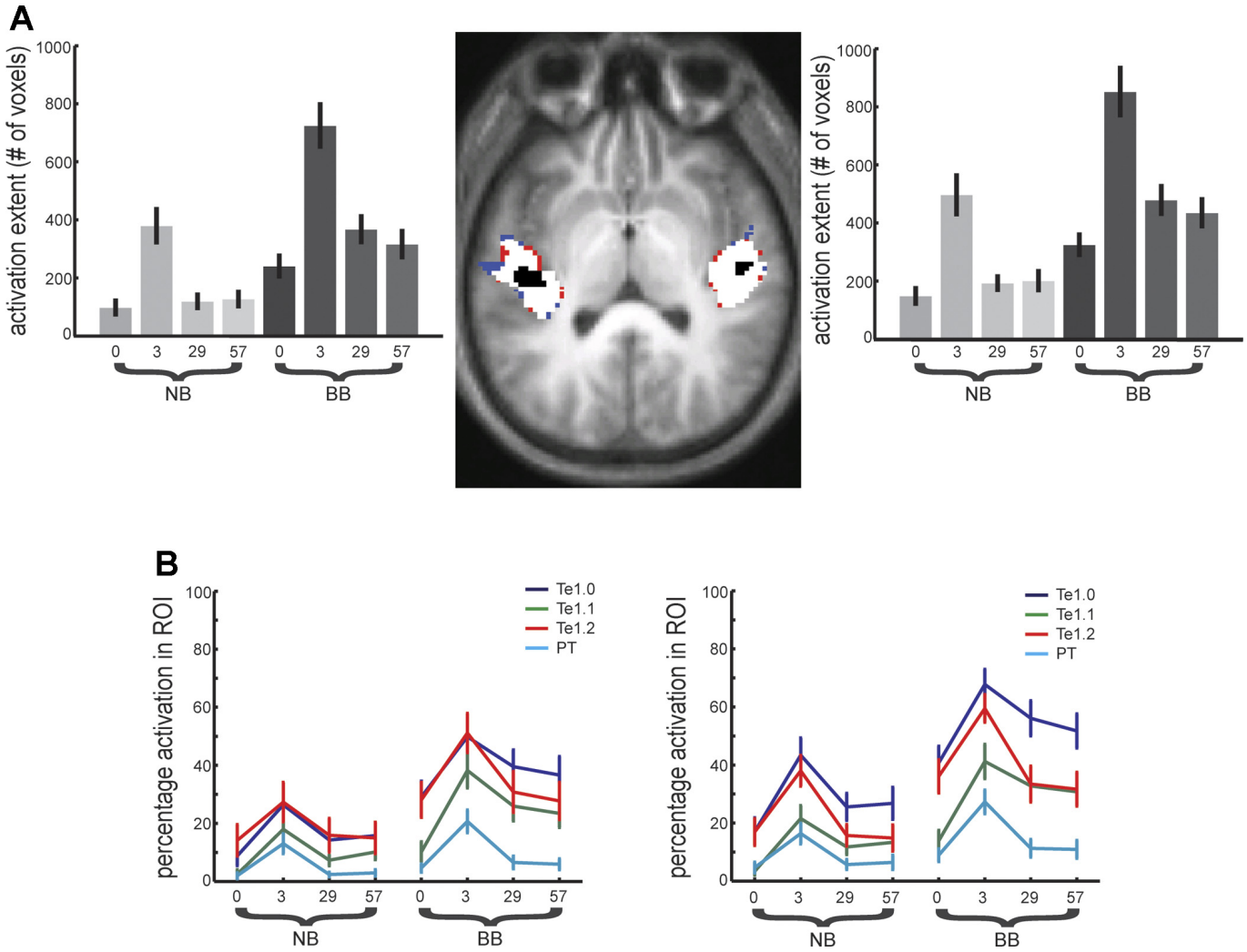


Fig. 5. Activation-extent analyses in *Study 2*. *A*: the bar plots display the overall mean activation extent (±SE) for the 8 experimental conditions in left and right auditory cortex, respectively. The overlay image in the middle depicts a visual combination of overall activation-extent and activation-strength analyses. Color code: white (NB and BB); blue (NB only); red (BB only); black (BB > NB). Note that in areas coded in blue (NB only) and red (BB only), the activation strengths between NB and BB are not statistically different from each other; only the areas in black reveal a statistically stronger response to BB than to NB sounds. Results are superimposed on a tilted (pitch: −0.5) axial section of participants' average structural scans. *B*: mean percentage activation extent (±SE) in the cortical ROIs Te1.0 (purple), Te1.1 (green), Te1.2 (red), and PT (cyan) in the 8 experimental conditions.

Temporal jittering of sAM

Previous studies investigating sensitivity to modulation rates have generally used constant rates (Giraud et al. 2000; Harms et al. 2005; Harms and Melcher 2002; Joris et al. 2004; Schönwiesner and Zatorre 2009). Natural sounds obviously do not have constant modulation rates but vary around mean modulation rates; for example, phonemes and syllables in speech have average rates of ~25–50 Hz and ~3–6 Hz, respectively (Rosen 1992). Therefore, the current stimulus using jittered sAM is arguably more ecologically valid than constant modulation rates, a claim that is supported by the greater sensitivity in the auditory system to these sounds compared with constant sAM stimuli.

Neuronal adaptation to a constant modulation rate is one possible explanation for these results (Bartlett and Wang 2005; Wojtczak and Viemeister 2005). Neurons within temporal modulation filters (Dau et al. 1997; Jepsen et al. 2008) would therefore habituate to the constant modulation rate, much like neuronal response characteristics in spectral filters. A second, complementary interpretation is that the jittered stimuli simply excited more modulation rate-sensitive neurons in total within a filter, since the AMs included not only one constant rate but also adjacent rates. The current design does not allow a definitive answer as to whether the greater signal for jittered stimuli reflects adaptation alone, recruitment of additional rate-sensitive neurons, or a combination of both, and further studies, e.g., using explicit auditory adaptation designs, are needed to dissociate these two possible mechanisms.

Spectral bandwidth

The results challenge the canonical hierarchical hypothesis that there exists a functional gradient between primary and nonprimary cortex with respect to spectral bandwidth (Chevillet et al. 2011; Rauschecker and Tian 2004; Rauschecker et al. 1995, 1997; Wessinger et al. 2001). In particular, the current study does not support the view that spectrally narrow sounds principally activate primary cortex

and that nonprimary areas do not respond vigorously to these sounds (but instead, only to spectrally broader sounds), as reported previously [see also Petkov et al. (2006) and Wessinger et al. (2001)].

The apparent difference with previous reports may therefore be due, at least in part, to the metrics used: whereas activation extent does increase with spectral bandwidth and thereby implicates the recruitment of later processing centers along the neuraxis, on a voxel-by-voxel basis, the effect of bandwidth, as assessed by activation strength, is only significant in core structures and is actually strongest in subcortical structures IC and MGB (Hawley et al. 2005). That is, for all of the ROIs tested, the effect size of bandwidth decreased from IC to primary cortex to nonprimary areas and was weakest in PT (in fact, in PT, only the 3-Hz sAM rates were significantly different between NB and BB sounds). This suggests that in nonprimary cortex, there is a considerable degree of subthreshold activity for NB sounds when comparing it against a silent baseline (as in the activation-extent metric), such that a direct comparison of BB and NB sounds (as in the activation-strength metric) does not reach statistical significance. This is in agreement with Harms et al. (2005; see page 218), who found that "response waveshape in posterior auditory cortex varied neither systematically nor substantially with large changes in bandwidth…". In fact, in their study, temporal modulation rate (and not bandwidth) was the main factor for differences in the hrf (see also the discussion of modulation rate below), and the companion paper by Wang et al. (2012) reports a similar pattern.

In the current study, the effect of spectral bandwidth is strongest in subcortical structures, and it is therefore not clear why it should be sharply attenuated in primary cortex and then reappear in nonprimary cortex (as the canonical hierarchical model would predict for cortex). One reason for this discrepancy might simply be sampling resolution: the cochleotopic representation is encompassed by a much greater neuronal volume in cortex compared with MGB and IC. From the study of Woods et al. (2010; e.g., see their Fig. 3), one can approximate that in human auditory cortex, 4 octaves span some 12 mm; that is, best-frequency doubles approximately every 3 mm, which would result in a critical band resolution of $\sim$1 mm. At the level of our voxel resolution ($2 \times 2 \times 2$ mm), this would indicate that each voxel encompasses approximately two critical bands. However, to our knowledge, this has not been tested explicitly. It should be noted, however, that the application of an 8-mm smoothing kernel (see MATERIALS AND METHODS) reduces the effective spatial resolution. Therefore, voxel-based metrics may not permit a straightforward comparison for activation strength (as opposed to neurophysiological recordings). For example, if a voxel contains several critical bands in the IC, then one would expect BB sounds (spanning 4 octaves) to elicit greater activation strengths than NB sounds (spanning 0.25 octave, i.e., within a critical band). Presumably, the integrated discharge rate across all neurons in the voxel would be greater for the BB sound. This is what we found in subcortical structures IC and MGB. In contrast, if the auditory cortex voxels encompass few critical bands, then the difference between BB and NB sounds is more likely to be associated with their relative efficacy.

In principle, a differential response to NB vs. BB sounds could be attributable to either of two factors: tonotopic organization or frequency-tuning properties of individual neurons. First, the decreasing effect of bandwidth from subcortical structures to nonprimary auditory cortex may be a function of the degree of tonotopic organization. For example, in mouse primary auditory cortex, the characteristic frequency at suprathreshold sound levels is quite variable for any given position along the tonotopic axis (Bandyopadhyay et al. 2010; Rothschild et al. 2010). Since tonotopy has also been shown to depend on the sound level at which tuning is assessed (Phillips et al. 1994), it is possible that a between-region comparison of spectral selectivity may depend on the sound level chosen for testing. Current in vivo measures from human auditory cortex generally do not address this issue [although, see Langers and van Dijk (in press)]. In fact, the precise pattern of tonotopic organization in human auditory cortex has been the subject of some debate (Schönwiesner et al. 2002; Talavage et al. 2000, 2004). Only recently, high-resolution, functional imaging studies are reaching a consensus by suggesting a tonotopic gradient orthogonal to HG (Da Costa et al. 2011; Langers and van Dijk, in press; Woods et al. 2010), which appears to extend, if only weakly, into nonprimary cortex (i.e., PT; c.f. Fig. 3 in Langers and van Dijkin, in press). Although human IC tonotopy is not well described, a recent nonhuman primate study demonstrated a robust frequency gradient (Baumann et al. 2011). Therefore, it is possible that the relative precision of tonotopy, from IC to PT, could explain our findings.

A second possibility is that the decreasing effect of bandwidth from subcortical structures to nonprimary auditory cortex may be a function of the strength of lateral or broadband inhibition. If so, then one might expect broadband sounds to suppress suprathreshold activity in regions with greater inhibition. In this case, the BOLD response would presumably depend both on inhibitory interneuron spiking and postsynaptic potentials (PSPs). Data from in vivo single neuron recordings show that suprathreshold tuning is regulated, in part, by inhibition (Foeller et al. 2001; Wang et al. 2002). Furthermore, frequency tuning of subthreshold stimuli is generally quite broad, with inhibitory and excitatory synaptic events contributing throughout the spectral range (Bandyopadhyay et al. 2010; Rothschild et al. 2010; Wehr and Zador 2003; Wu et al. 2008). Theoretical studies based on large-scale, biologically plausible models of primary auditory cortex suggest that the precise configuration of synaptic inhibition can lead to dramatic changes to the input-output characteristics (Levy and Reyes 2011). To our knowledge, there exists no information about the relative strength of lateral or co-tuned inhibition in human auditory cortex. Therefore, the impact of differences between local network properties on our findings cannot presently be evaluated.

A recent study (Chevillet et al. 2011) also addressed the question of an anatomical hierarchy for spectral bandwidth or spectral complexity and is therefore directly relevant to our study. The authors compared activation in human auditory cortex for pure tones, bandpassed (1 octave), and broadband white noise, as well as vowel sounds. They found increasing activation extents as a function of spectral bandwidth and spectral complexity (vowel sounds). Chevillet and colleagues (2011; see their Fig. 2) also apply a binary decision to the activation-strength metric; the results are in agreement with our results, in that they show significantly stronger responses to bandpassed and broadband noise compared with pure tones in

primary cortex (and not in nonprimary cortex). The authors (Chevillet et al. 2011; see page 9347) interpret this as indicating "that more neurons in core, each responding to select frequencies contained in the BPN [bandpassed noise], are recruited by the spectrally wider BPN". However, this would contradict the original finding that neurons in primary cortex do not respond to broadband noise (Rauschecker et al. 1995), and it remains unclear why—if there exists a gradient with respect to spectral bandwidth—nonprimary cortex did not show the clearest pattern of stronger responses to broad band compared with NB (or pure tone) sounds.

For the detailed analysis of activation extent within ROIs, the hierarchical framework would have predicted an interaction, such that NB sounds lead to greater activation extents in primary rather than in nonprimary cortex, whereas BB sounds would elicit the reverse pattern. This is not what we found; BB sounds, in general, led to greater activation extents in all ROIs (cf. main effects in Table 2, Percentage Activation Extent, and the significant pairwise comparisons). That is, the percentage activation extent analysis partly confirms the overall activation extent results, in that BB sounds lead to larger activations within each ROI, with the caveat that there is no evidence for an interaction along the lines of preference for NB sounds in primary cortex.

The present findings, suggesting a greater response to BB stimuli at the level of the IC and perhaps the MGB, may reflect the limited spatial resolution of the analytical procedures discussed above and the reliance on the activation-strength metric. Neuronal discharge rate in IC is much higher than in auditory cortex of awake animals, particularly when response adaptation is taken into account (Ter-Mikaelian et al. 2007). Furthermore, although IC neuron responses to stimulus spectrum are diverse, the average discharge rate elicited by pulsed tones at characteristic frequency is similar to or greater than that evoked by BB noise (Aitkin 1991; Aitkin et al. 1994; Ehret and Merzenich 1988; Recanzone 2000). Therefore, it seems likely that the greater responses to BB sounds simply reflect the recruitment of a larger fraction of neurons within the few sampling voxels.

At the same time, NB sounds (spanning 0.25 octave) were centered on different frequencies; it is therefore conceivable that the response to these sounds adapted less, simply because roving the center frequency resulted in longer durations between repetitions of that particular NB and center-frequency combination. This could theoretically lead to an increased response to NB sounds that might equal that of BB sounds (whose spectral region is the same throughout). The alternative approach would have been to keep the center frequency the same throughout for NB sounds; however, in that case, any difference between BB and NB sounds could be attributed to the fact that BB sounds covered a significantly different spectral region throughout. Previous studies, with which we aim to compare [e.g., Chevillet et al. (2011), Harms et al. (2005), and Wessinger et al. (2001)], have used a similar approach of roving the center frequency.

One note of caution concerns the inferences that one can draw when comparing different techniques, such as electrophysiology in nonhuman primates and fMRI in humans. Electrophysiological recordings are based on a sparse but spatially precise sampling of neurons in a restricted cortical (or subcortical) area, whereas the BOLD contrast in fMRI is based on the hemodynamic activity evoked by comparatively vast ensembles of neurons [e.g., Logothetis (2008)]. Furthermore, many previous studies that indicated a preference for broadband sounds in association areas used neuronal discharge rates or multi-unit activity (MUA) as the dependent variable, whereas the BOLD signal is more closely correlated with synaptic activity, e.g., local field potentials (LFPs) (Arthurs and Boniface 2002; Goense and Logothetis 2008; Logothetis et al. 2001). LFP and MUA measures can also be distinguished via their spatial and temporal characteristics: LFP activity integrates over an area of $\sim$2 mm^2 (Malonek and Grinvald 1996), whereas MUA activity integrates over a much smaller region, on the order of 200 μm^2 (Thompson et al. 2003); similarly, LFPs or PSPs extend over much longer time scales than the very brief bursts of action potentials. These characteristics might explain why the BOLD signal corresponds more closely to LFP than to MUA measures.

Furthermore, the increased stimulus selectivity observed with single-unit recordings in auditory cortex (Wang 2007) may simultaneously be associated with substantial subthreshold synaptic activity to noneffective stimuli. Thus it is conceivable that a summation of a fair amount of subthreshold or LFP activity is responsible for the larger (both strength and extent) fMRI responses to BB stimuli, whereas an assessment of discharge rate would display just the opposite relationship. It is also possible that the increased response to BB compared with NB sounds (or pure tones; see Table 1) seen in fMRI reflects the fact that BB sounds would activate a broader population of cells (although, perhaps be associated with less-vigorous neuronal firing), and this may not be apparent from sparse sampling in electrophysiological recordings. Consequently, it is possible that these two techniques could produce seemingly different data that, however, do not necessarily contradict each other.

Finally, spectral bandwidth as such may, after all, not be an ideal metric to describe hierarchical processing in the auditory system but rather, spectral complexity. Neither white noise (as used in most previous studies) nor pink noise (as used here) displays complexly structured organization (rather, they are completely random), and it is possible that association areas would indeed show a preference for such complexity [e.g., see Chevillet et al. (2011)], whereas white/pink noise is "uninteresting" because of its lack of structure.

Temporal modulation rate

The data illustrate a striking sensitivity to slow modulation rates around 3 Hz, which arises in cortex and increases from primary to association areas, consistent with data from other labs, which have also tested other slow modulation rates (Giraud et al. 2000; Harms et al. 2005; Harms and Melcher 2002). Importantly, this appears to be a bandpass- and not a lowpass-filter effect, as it does not include the static sounds (0 Hz sAM) and thus indicates the perceptual importance of slow modulation rates within a restricted band.

Previous imaging studies investigating the representation of temporal modulation rates in the auditory system suggest: *1*) a preference for increasingly slow temporal modulation rates as one ascends through the auditory system (Giraud et al. 2000), and *2*) increasingly phasic, hemodynamic response waveshape characteristics for faster modulation rates (Giraud et al. 2000;

Harms et al. 2005; Harms and Melcher 2002). In their seminal studies, Harms and colleagues (2005) and Harms and Melcher (2002, 2003) used similar stimulus parameters as in the current study (e.g., Sin, NB, and BB sounds presented at 2 Hz or 35 Hz modulation rates) and showed that modulation rate (but not spectral bandwidth) was the main determinant of response waveshape type (e.g., sustained vs. phasic).

The current study extends these findings in three ways. First, we show that these results also hold for jittered modulation rates and are therefore less likely an artifact of laboratory sounds (e.g., constant sAM); similarly, the use of pink noise instead of white noise adds to that conclusion, since many naturally occurring sounds display 1/f power spectra (Voss and Clarke 1975), and sounds derived from such spectra have been successful in elucidating principles of auditory processing (Overath et al. 2007; Patel and Balaban 2000; Schmuckler and Gilden 1993). Second, the increased spatial resolution (e.g., $2 \times 2 \times 2$ vs. $3.1 \times 3.1 \times 7$ mm) and coverage (e.g., much of the auditory system starting with IC vs. single slice) provide an improvement of spatial description of temporal modulation rate and spectral bandwidth processes in the auditory system. Third, the current study uses robust random-effects statistics at conservative significance thresholds, thereby solidifying the value and scope of the previous results from a series of individual case studies to the level of population inference.

Another aspect concerns the spatial organization of modulation filters. There is, as yet, no clear evidence for an orderly representation of temporal modulation rate filters (Dau et al. 1997; Jepsen et al. 2008) in auditory cortex (Giraud et al. 2000; Schönwiesner and Zatorre 2009). Our data also did not reveal any systematic representation of modulation rates in auditory cortex (apart from the strong sensitivity to slow modulation rates), and a much more fine-grained coverage of modulation rates, as well as possibly different types of designs (e.g., fMRI adaptation designs), is needed to address this issue thoroughly.

One unexpected finding concerns the absence of an effect of temporal modulation rate in IC and MGB. It should be mentioned, however, that the modulation rates used here may not be optimal for subcortical structures, such as IC, which based on neurophysiological recordings in a variety of species, is more sensitive to slightly faster rates (Joris et al. 2004; e.g., see their Fig. 9). It is therefore conceivable that the inclusion of higher modulation rates would have revealed a main effect of modulation rate (i.e., a greater sensitivity to faster than to slower modulation rates). Giraud et al. (2000) and Harms et al. (2002) report such a gradient [in the case of Giraud and colleagues (2000), the modulation rates ranged well into the realm of pitch with 256 Hz sAM], and it may be that jittering the modulation rate, as in the current study, perturbed phase-locking in subcortical structures. Incidentally, IC was the only area that showed no main effect of modulation type for activation strength in *Study 1*. This may be due to the broader modulation transfer functions (MTFs) in IC compared with subsequent structures, since MTFs generally become narrower as one ascends the auditory system (Joris et al. 2004).

Finally, the companion paper (Wang et al. 2012) also suggests a (monotonic) relationship between modulation rate and response amplitude, as indexed by the power of the phase-locked (or stimulus-synchronized) MEG response. Such complementary evidence from a different methodology strengthens the conclusion that at the level of cortex, modulation rate is an important index of auditory processing principles.

Interaction between bandwidth and modulation rate

A central aim of this study was to determine whether there is evidence for an interaction between spectral bandwidth and temporal modulation rate in the auditory system. According to the uncertainty principle of neural coding (Zatorre and Belin 2001; Zatorre et al. 2002), high spectral resolution implies narrow bandwidth tuning at the expense of temporal resolution (i.e., slow temporal modulation rates), whereas high temporal resolution implies fast modulation rate tuning at the expense of spectral resolution (i.e., broad spectral filters). It has also been suggested that faster modulation rates may be more left lateralized, whereas there is growing consensus that slower modulations are right lateralized (Boemio et al. 2005; Overath et al. 2008; Schönwiesner et al. 2005; Zatorre and Belin 2001). We did not find any evidence for such an interaction of spectral bandwidth and temporal modulation rate; in fact, with the imaging and stimulus parameters used here, no subcortical or cortical area showed an increase in signal strength or activation extent as a function of increasing modulation rate or decreasing spectral bandwidth. This would indicate that spectral bandwidth and temporal modulation rate are processed independently. At the same time, it can be argued (Joris et al. 2004) that a trade-off between temporal modulation rate and spectral bandwidth only comes into effect when the spectral bandwidth is narrower than the value of the modulation rate, i.e., when the filter transmitting the modulation rate becomes too narrow to transmit the bandwidth of the modulation. In the present case, with AM rates below 60 Hz, this was not the case, which might explain the absence of an interaction in our results.

Conclusion

Our results highlight a distinct pattern along the auditory system with respect to processing spectral bandwidth and temporal modulation rates, whereby sensitivity to spectral bandwidth decreases, and sensitivity to slow modulation rates increases. Second, temporally jittered sAM leads to stronger responses than constant sAM, possibly by mitigating against stimulus-specific adaptation. Finally, the present results suggest that putative parallels between visual- and auditory-processing principles are not necessarily straightforward. For example, whereas the differentiation between "simple" and "complex" cells may be useful in the visual system, there is less-convincing evidence for such a distinction, or on what it would be based, in the auditory system. In fact, it has been argued that IC and the primary visual cortex (V1) occupy an equivalent level of signal processing in the auditory and visual hierarchies, respectively, which would put signal processes in primary auditory cortex at the level of approximately area V4 in the visual system (King and Nelken 2009).

ACKNOWLEDGMENTS

We thank Adrian Rees for helpful comments on an earlier version of this manuscript.

GRANTS

Support for this study was provided by the National Institute on Deafness and Other Communication Disorders Grant 2R01DC05660 to D. Poeppel.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.O., D.H.S., and D.P. conception and design of research; T.O. and Y.Z. performed experiments; T.O. analyzed data; T.O., D.H.S., and D.P. interpreted results of experiments; T.O. prepared figures; T.O. drafted manuscript; T.O., D.H.S., and D.P. edited and revised manuscript; T.O., Y.Z., D.H.S., and D.P. approved final version of manuscript.

REFERENCES

Aitkin L. Rate-level functions of neurons in the inferior colliculs of cats measured with the use of free-field sound stimuli. *J Neurophysiol* 65: 383–392, 1991.

Aitkin L, Tran L, Syka J. The responses of neurons in subdivisions of the inferior colliculus of cats to tonal, noise and vocal stimuli. *Exp Brain Res* 98: 53–64, 1994.

Anderson LA, Christianson GB, Linden JF. Stimulus-specific adaptation occurs in the auditory thalamus. *J Neurosci* 29: 7359–7363, 2009.

Arthurs OJ, Boniface S. How well do we understand the neural origins of the fMRI BOLD signal? *Trends Neurosci* 25: 27–31, 2002.

Atencio CA, Schreiner CE. Spectrotemporal processing differences between auditory cortical fast-spiking and regular-spiking neurons. *J Neurosci* 28: 3897–3910, 2008.

Bandyopadhyay S, Shamma SA, Kanold PO. Dichotomy of functional organization in the mouse auditory cortex. *Nat Neurosci* 13: 361–368, 2010.

Bartlett EL, Wang X. Neural representations of temporally-modulated signals in the auditory thalamus of awake primates. *J Neurophysiol* 97: 1005–1117, 2007.

Bartlett EL, Wang XQ. Long-lasting modulation by stimulus context in primate auditory cortex. *J Neurophysiol* 94: 83–104, 2005.

Baumann S, Griffiths TD, Rees A, Hunter D, Sun L, Thiele A. Characterisation of the BOLD response time course at different levels of the auditory pathway in non-human primates. *Neuroimage* 50: 1099–1108, 2010.

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

Bendor D, Wang X. Differential neural coding of acoustic flutter within primate auditory cortex. *Nat Neurosci* 10: 763–771, 2007.

Boemio A, Fromm S, Braun A, Poeppel D. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci* 8: 389–395, 2005.

Brett M, Anton, J-L, Valabregue R, Poline J-B. Region of interest analysis using an SPM toolbox (abstract). *NeuroImage* 16: 2002.

Cardin JA, Palmer LA, Contreras D. Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex. *J Neurosci* 27: 10333–10344, 2007.

Carrasco A, Lomber SG. Differential modulatory influences between primary auditory cortex and the anterior auditory field. *J Neurosci* 29: 8350–8362, 2009a.

Carrasco A, Lomber SG. Evidence for hierarchical processing in cat auditory cortex: nonreciprocal influence of primary auditory cortex on the posterior auditory field. *J Neurosci* 29: 14323–14333, 2009b.

Chevillet M, Riesenhuber M, Rauschecker JP. Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J Neurosci* 31: 9345–9352, 2011.

Cohen J. Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educ Psychol Meas* 33: 107–112, 1973.

Da Costa S, van der Zwaag W, Marques JP, Frackowiak RS, Clarke S, Saenz M. Human primary auditory cortex follows the shape of Heschl's gyrus. *J Neurosci* 31: 14067–14075, 2011.

Dau T, Kollmeier B, Kohlrausch A. Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J Acoust Soc Am* 102: 2892–2905, 1997.

Devlin JT, Sillery EL, Hall DA, Hobden P, Behrens TEJ, Nunes RG, Clare S, Matthews PM, Moore DR, Johansen-Berg H. Reliable identification of the auditory thalamus using multi-modal structural analyses. *Neuroimage* 30: 1112–1120, 2006.

Ehret G, Merzenich MM. Neuronal discharge rate is unsuitable for encoding sound intensity at inferior-colliculs level. *Hear Res* 35: 1–7, 1988.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25: 1325–1335, 2005.

Elliott TM, Theunissen FE. The modulation transfer function for speech intelligibility. *PLoS Comput Biol* 5: e1000302, 2009.

Foeller E, Vater M, Kössl M. Laminar analysis of inhibition in the gerbil primary auditory cortex. *J Assoc Res Otolaryngol* 2: 279–296, 2001.

Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2: 189–210, 1995.

Giraud AL, Lorenzi C, Ashburner J, Wable J, Johnsrude I, Frackowiak R, Kleinschmidt A. Representation of the temporal envelope of sounds in the human brain. *J Neurophysiol* 84: 1588–1598, 2000.

Goense JD, Logothetis NK. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr Biol* 18: 631–640, 2008.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. "Sparse" temporal sampling in auditory fMRI. *Hum Brain Mapp* 7: 213–223, 1999.

Harms MP, Guinan JJ, Sigalovsky IS, Melcher JR. Short-term sound temporal envelope characteristics determine multisecond time patterns of activity in human auditory cortex as shown by fMRI. *J Neurophysiol* 93: 210–222, 2005.

Harms MP, Melcher JR. Detection and quantification of a wide range of fMRI temporal repsonses using a physiologically-motivated basis set. *Hum Brain Mapp* 20: 168–183, 2003.

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

Hawley ML, Melcher JR, Fullerton BC. Effects of sound bandwidth on fMRI activation in human auditory brainstem nuclei. *Hear Res* 204: 101–110, 2005.

Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol (Paris)* 195: 215–243, 1968.

Jepsen ML, Ewert SD, Dau T. A computational model of human auditory signal processing and perception. *J Acoust Soc Am* 124: 422–438, 2008.

Joris PX, Schreiner CE, Rees A. Neural processing of amplitude-modulated sounds. *Physiol Rev* 84: 541–577, 2004.

King AJ, Nelken I. Unraveling the principles of auditory cortical processing: can we learn from the visual system? *Nat Neurosci* 12: 698–701, 2009.

Langers DR, van Dijk P. Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. *Cereb Cortex*. In press.

Langers DR, van Dijk P, Schoenmaker ES, Backes WH. fMRI activation in relation to sound intensity and loudness. *Neuroimage* 35: 709–718, 2007.

Levy RB, Reyes AD. Coexistence of lateral and and co-tuned inhibitory configurations in cortical networks. *PLoS ONE* 7: e1002161, 2011.

Logothetis NK. The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos Trans R Soc Lond B Biol Sci* 357: 1003–1037, 2002.

Logothetis NK. What we can do and what we cannot do with fMRI. *Nature* 453: 869–878, 2008.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412: 150–157, 2001.

Lu T, Liang L, Wang X. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 4: 1131–1138, 2001.

Ma WP, Liu BH, Li YT, Huang ZJ, Zhang LI, Tao HW. Visual representations by cortical somatostatin inhibitory neurons—selective but with weak and delayed responses. *J Neurosci* 30: 14371–14319, 2010.

Malmierca MS, Cristaudo S, Pérez-González D, Covey E. Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29: 5483–5493, 2009.

Malone BJ, Scott BH, Semple MN. Context-dependent adaptive coding of interaural phase disparity in the auditory cortex of awake macaques. *J Neurosci* 22: 4625–4638, 2002.

Malone BJ, Semple MN. Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. *J Neurophysiol* 86: 1113–1130, 2001.

Malonek D, Grinvald A. Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. *Science* 272: 551–554, 1996.

Maunsell JH, Newsome WT. Visual processing in monkey extrastriate cortex. *Annu Rev Neurosci* 10: 363–401, 1987.

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

Overath T, Cusack R, Kumar S, von Kriegstein K, Warren JD, Grube M, Carlyon RP, Griffiths TD. An information theoretic characterisation of auditory encoding. *PLoS Biol* 5: e288, 2007.

Overath T, Kumar S, von Kriegstein K, Griffiths TD. Encoding of spectral correlation over time in auditory cortex. *J Neurosci* 28: 13268–13273, 2008.

Pandya PK, Rathbun DL, Moucha R, Engineer ND, Kilgard MP. Spectral and temporal processing in rat posterior auditory cortex. *Cereb Cortex* 18: 301–314, 2009.

Patel AD, Balaban E. Temporal patterns of human cortical activity reflect tone sequence structure. *Nature* 404: 80–84, 2000.

Petkov CI, Kayser C, Augath M, Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol* 4: e215, 2006.

Phillips DP, Semple MN, Calford MB, Kitzes LM. Level-dependent representation of stimulus frequency in cat primary auditory cortex. *Exp Brain Res* 102: 210–226, 1994.

Pickett JM. *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology*. Boston, MA: Allyn & Bacon, 1999.

Pierce CA, Block RA, Aguinis H. Cautionary note on reporting eta-squared values from multifactor ANOVA designs. *Educ Psychol Meas* 64: 916–924, 2004.

Pressnitzer D, Patterson RD, Krumbholz K. The lower limit of melodic pitch. *J Acoust Soc Am* 109: 2074–2084, 2001.

Rauschecker JP, Tian B. Processing of band-passed noise in the lateral auditory belt of the rhesus monkey. *J Neurophysiol* 91: 2578–2598, 2004.

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

Rauschecker JP, Tian B, Pons T, Mishkin M. Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol* 382: 89–103, 1997.

Recanzone GH. Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. *Hear Res* 150: 104–118, 2000.

Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336: 367–373, 1992.

Rothschild G, Nelken I, Mizrahi A. Functional organization and population dynamics in the mouse primary auditory cortex. *Nat Neurosci* 13: 353–360, 2010.

Schmuckler MA, Gilden DL. Auditory perception of fractal contours. *J Exp Psychol Hum Percept Perform* 19: 641–660, 1993.

Schönwiesner M, Rübsamen R, von Cramon DY. Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur J Neurosci* 22: 1521–1528, 2005.

Schönwiesner M, von Cramon DY, Rübsamen R. Is it tonotopy after all? *Neuroimage* 17: 1144–1161, 2002.

Schönwiesner M, Zatorre RJ. Spectro-temporal transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. *Proc Natl Acad Sci USA* 106: 14611–14616, 2009.

Simons DJ, Carvell GE. Thalamocortical response transformation in the rat vibrissa/barrel system. *J Neurophysiol* 61: 311–330, 1989.

Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 114: 3394–3411, 2003.

Talavage TM, Ledden PJ, Benson RR, Rosen BR, Melcher JR. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. *Hear Res* 150: 225–244, 2000.

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

Tan AY, Zhang LI, Merzenich MM, Schreiner CE. Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. *J Neurophysiol* 92: 630–643, 2004.

Ter-Mikaelian M, Sanes DH, Semple MN. Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. *J Neurosci* 27: 6091–6102, 2007.

Thompson JK, Peterson MR, Freeman RD. Single-neuron activity and tissue oxygenation in the cerebral cortex. *Science* 299: 1070–1072, 2003.

Ulanovsky N, Las L, Farkas D, Nelken I. Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24: 10440–10453, 2004.

Viswanathan A, Freeman RD. Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. *Nat Neurosci* 10: 1308–1312, 2007.

Voss RF, Clarke J. 1/f noise in music and speech. *Nature* 258: 317–318, 1975.

Wang J, McFadden SL, Caspary D, Salvi R. Gamma-aminobutyric acid circuits shape response properties of auditory cortex neurons. *Brain Res* 944: 219–231, 2002.

Wang X. Neural coding strategies in auditory cortex. *Hear Res* 229: 81–93, 2007.

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ. Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *J Neurophysiol* 107: 2033–2041, 2012.

Wehr M, Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426: 442–446, 2003.

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

Westbury CF, Zatorre RJ, Evans AC. Quantifying variability in the planum temporale: a probability map. *Cereb Cortex* 9: 392–405, 1999.

Wojtczak M, Viemeister NF. Forward masking of amplitude modulation: basic characteristics. *J Acoust Soc Am* 118: 3198–3210, 2005.

Woods DL, Herron TJ, Cate AD, Yund EW, Stecker GC, Rinne T, Kang X. Functional properties of human auditory cortical fields. *Front Syst Neurosci* 3: 155, 2010.

Wu GK, Arbuckle R, Liu BH, Zhang LI. Lateral sharpening of cortical frequency tuning by approximately balanced inhibition. *Neuron* 58: 132–143, 2008.

Xie R, Gittelman JX, DPG. Rethinking tuning: in vivo whole-cell recordings of the inferior colliculus in awake bats. *J Neurosci* 27: 9469–9481, 2007.

Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11: 946–953, 2001.

Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6: 37–46, 2002.

Zhang Y, Suga N. Modulation of responses and frequency tuning of thalamic and collicular neurons by cortical activation in mustached bats. *J Neurophysiol* 84: 325–333, 2000.

Zhang Z, Liu CH, Yu YQ, Fujimoto K, Chan YS, He J. Corticofugal projection inhibits the auditory thalamus through the thalamic reticular nucleus. *J Neurophysiol* 99: 2938–2945, 2008.

Zhou X, Jen PH. Brief and short-term corticofugal modulation of subcortical auditory responses in the big brown bat, *Eptesicus fuscus*. *J Neurophysiol* 84: 3083–3087, 2000.