

# FM-selective Networks in Human Auditory Cortex Revealed Using fMRI and Multivariate Pattern Classification

I-Hui Hsieh<sup>1\*</sup>, Paul Fillmore<sup>2\*</sup>, Feng Rong<sup>2\*</sup>, Gregory Hickok<sup>2</sup>,  
and Kourosh Saberi<sup>2</sup>

## Abstract

■ Frequency modulation (FM) is an acoustic feature of nearly all complex sounds. Directional FM sweeps are especially pervasive in speech, music, animal vocalizations, and other natural sounds. Although the existence of FM-selective cells in the auditory cortex of animals has been documented, evidence in humans remains equivocal. Here we used multivariate pattern analysis to identify cortical selectivity for direction of a multitone FM sweep. This method distinguishes one pattern of neural activity from another within the same ROI, even when overall level of activity is similar, allowing for di-

rect identification of FM-specialized networks. Standard contrast analysis showed that despite robust activity in auditory cortex, no clusters of activity were associated with up versus down sweeps. Multivariate pattern analysis classification, however, identified two brain regions as selective for FM direction, the right primary auditory cortex on the supratemporal plane and the left anterior region of the superior temporal gyrus. These findings are the first to directly demonstrate existence of FM direction selectivity in the human auditory cortex. ■

## INTRODUCTION

Frequency modulation (FM) is a basic acoustic component of all complex sounds from speech and music to animal vocalizations in mammals, marine species, birds, and even insect acoustics (Sabourin, Gottlieb, & Pollack, 2008; Dankiewicz, Helweg, Moore, & Zafran, 2002; Dear, Simmons, & Fritz, 1993; Klump & Langemann, 1992; Coscia, Phillips, & Fentress, 1991; Ryan & Wilczynskin, 1988; Huber & Thorson, 1985; Fant, 1970). In human speech, frequency glides and formant transitions provide critical cues to phonemic identification (Divenyi, 2009; Gordon & O'Neill, 1998; Pickett, 1980; Fant, 1970; Liberman, Delattre, Gerstman, & Cooper, 1956) and, additionally, in tonal languages (e.g., Mandarin or Thai) play an important role in lexical distinction (Luo, Wang, Poeppel, & Simon, 2007; Stagger, Downs, & Sommers, 1992; Howie, 1976). FM sweeps have also been shown to influence a wide range of clinical-translational and perceptual phenomena from language-based learning impairments (Subramanian, Yairi, & Amir, 2003; Tallal et al., 1996) to electric hearing in cochlear implant patients (Chen & Zeng, 2004), auditory object formation (Carlyon, 1994), and music perception (d'Alessandro, Rosset, & Rossi, 1998).

Animal neurophysiological studies have identified populations of FM-selective neurons in brainstem structures (e.g., the inferior colliculus) and higher levels of the auditory cortex (Razak & Fuzessery, 2006, 2010; Williams & Fuzessery, 2010; Gittelman, Li, & Pollak, 2009; Kajikawa et al., 2008; Andoni, Li, & Pollak, 2007; Godey, Atencio, Bonham, Schreiner, & Cheung, 2005; Woolley & Casseday, 2005; Koch & Grothe, 1998; Fuzessery & Hall, 1996; Suga, 1968). Neuroimaging studies of human cortical discrimination of FM sweeps have identified brain regions with increased activity during discrimination of sweep direction or during passive listening to FM glides by contrasting activity in these regions to activation levels either at rest or during performance of other auditory tasks (e.g., categorical perception of CV syllables or word/nonword lexical decisions). In general, these studies implicate the right auditory cortex during identification of sweep direction, especially for slower rate FMs, and either bilaterally or the left hemisphere during tasks involving discrimination of sweep duration particularly for stimuli characterized by faster sweeps (Behne, Scheich, & Brechmann, 2005; Brechmann & Scheich, 2005; Poeppel et al., 2004; Hall et al., 2002; Binder et al., 2000; Thivard, Belin, Zilbovicius, Poline, & Samson, 2000; Belin et al., 1998; Scheich et al., 1998; Schlosser, Aoyagi, Fulbright, Gore, & McCarthy, 1998; Johnsrude, Zatorre, Milner, & Evans, 1997). Consistent with these findings, human nonaphasic patients with lesions to their right cortical hemisphere as well as animals

<sup>1</sup>National Central University, Jhongli City, Taiwan, <sup>2</sup>University of California, Irvine

\*Contributed equally to this work.

with lesions to the right auditory cortex display a significant decline in discrimination of FM sweep direction (Wetzel, Ohl, Wagner, & Scheich, 1998; Divenyi & Robinson, 1989).

The current study was motivated by three specific considerations. First, all prior human neuroimaging studies of FM coding, to our knowledge, have sought to identify brain regions recruited for task classification (e.g., FM coding contrasted to CV classification) and not regions selective for a within stimulus class feature (e.g., sweep direction; Brechmann & Scheich, 2005). The FM regions identified by these studies have been shown to be more active during behavioral identification or discrimination of sweep direction when contrasted to activity associated with other non-FM tasks or with rest (no stimulus). No within-stimulus class distinction is made between features such as up versus down sweeps. This is a critical point because the identified regions are not derived from pooling cortical responses to those trials on which an up sweep is presented, contrasted to those on which a down sweep occurs, and therefore, cannot conclusively establish existence of networks in the human cortex selective for FM sweep direction or rate. At most, these studies have only identified regions active during one perceptual/cognitive decision versus another—regions that may potentially also be recruited for other pitch-related tasks or decisions.

Second, if direction-selective FM neurons do in fact exist in the human cortex, it is likely, as suggested by animal neurophysiology (Tian & Rauschecker, 2004), that they are interspersed within the same general cortical regions, and therefore, unless there is a significantly larger number of units that classify one sweep type (e.g., up sweep), contrast methods used by all prior neuroimaging studies of FM coding would not reveal existence of such putative direction-selective neurons. It is therefore not surprising that, given these limitations of contrast-based methods, no prior neuroimaging study of FM coding has reported on differences in activity patterns associated with different features of FM sounds (hence, verify existence of such units). In the current study, we used multivariate pattern classification analysis (MVPA) to identify brain regions selectively responsive to sweep direction. This method can distinguish one pattern of neural activity from another within the same ROI and hence determine whether and where there may exist networks of neurons selective for features of FM sounds.

Third, most prior neuroimaging, neurophysiological, and psychophysical studies of FM sweep coding have used single tones as stimuli. Speech, music, and other band-limited modulated sounds are highly complex signals with spectrotemporally dynamic multicomponent structures. The FM components of speech sounds such as formant transitions, for example, are embedded within composite acoustic structures and are not typically perceived as separate entities or identified as isolated sweeps in pitch. However, simply changing the direction of a formant transition sweep within the complex from up to down can categorically change the percept of one voiced sound to another

(e.g., from “ba” to “ga”; Liberman, Harris, Hoffman, & Griffith, 1957; Delattre, Liberman, & Cooper, 1955). We know of only three prior neuroimaging studies that have employed multitone structures as FM stimuli, with two using linear sweeps (Behne et al., 2005; Brechmann & Scheich, 2005) and one using a sinusoidal modulator (Hall et al., 2002). None have investigated differences in cortical activity patterns based on contrasting features of the FM sound (e.g., up vs. down sweep). Furthermore, interpretation of findings from these studies is complicated by existence of potential spectral edge pitch artifacts, that is, a shift in the position of the FM band, and hence a shift in the centroid of spectral energy across the start and end points of stimulation. A narrowband filter centered on the terminal sweep frequency may be used by the system as a simple energy detector to successfully perform a direction-discrimination task without requiring the coding of frequency sweep per se. The current study employed a multitone FM complex bounded by fixed-frequency tones, designed specifically to maintain constant bandwidth and band position throughout stimulus duration regardless of sweep direction. We found that, although traditional contrast methods did not, and likely could not, reveal selectivity for FM sweep direction, multivariate pattern classification identified two regions, the left anterior and right primary areas of the superior temporal gyrus (STG) selective for sweep direction.

## METHODS

### Participants

Twenty-two right-handed speakers of Mandarin Chinese (11 women) between 19 and 30 years (mean = 23.4) participated in the study. One volunteer was excluded due to poor data quality. All volunteers were students at the National Central University (Taiwan) and had normal hearing, no known history of neurological disease, and no other contraindications for MRI. Written informed consent was obtained from each participant in accordance with the National Central University Institutional Review Board guidelines.

### Stimuli

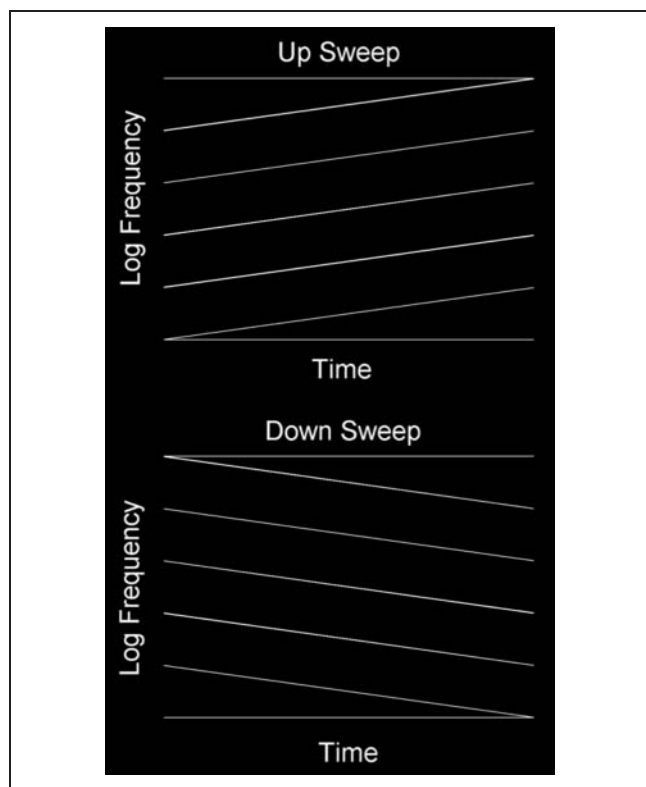
Stimuli were generated digitally using Matlab software (Mathworks, Inc., Natick, MA) on a Dell PC (Optiplex GX270) and presented diotically at a rate of 44.1 kHz through 16-bit digital-to-analog converters and through MR compatible headphones (Sensimetrics S14, Malden, MA). Stimuli consisted of five-tone logarithmic FM sweep complexes generated from Equation 1 (Hsieh & Saberi, 2009),

$$X(t) = \sum_{n=1}^5 \sin \left( 2\pi \frac{f_s(n)}{\ln(f_e(n)/f_s(n))} \left[ \exp \left( \frac{\ln(f_e(n)/f_s(n))}{T_s} t \right) - 1 \right] \right) \quad (1)$$

where  $f_s(n)$  and  $f_e(n)$  represent the starting and ending sweep frequencies of the  $n$ th tone (in hertz) and  $T_s$  is stimulus duration (in seconds). Schematic diagrams of up and down sweep complexes are shown in Figure 1. There were four experimental conditions ( $2 \times 2$  design): two sweep directions (up or down in frequency) at two stimulus durations ( $T_s = 100$  or  $400$  msec corresponding to FM rates of 3.3 and 0.83 octaves/sec, respectively). The only purpose of using two durations/rates (slow and fast) was to maximize the likelihood of observing a sweep direction effect, the main variable of interest in our study. This was done because neurophysiological studies in animals have shown that FM cells are not only direction selective but also rate selective, and it was a priori unknown to what rates such putative cells in human auditory cortex might be tuned. All FM complexes had a base starting frequency of 500 Hz (i.e., lowest frequency component), with the remaining four FM tones having starting frequencies that were spaced  $1/3$  octave apart ( $f_s(n) = f_s(n-1) \cdot 2^{1/3}$  for  $n > 1$ ), resulting in starting frequencies of 500, 630, 794, 1000, and 1260 Hz, respectively, for the five tones of the complex. Each FM tone also swept through  $1/3$  of an octave for a given stimulus duration, resulting in end

frequencies that matched exactly the start frequency of the next higher tone with  $f_e(n) = 630, 794, 1000, 1260$ , and  $1587$  Hz (e.g.,  $\log_2(630/500) = 0.333$ ). We chose logarithmic spacing between components because cortical tonotopy is logarithmically organized, with a compressive spacing at higher frequencies (Pickles, 2008). Use of log spacing between components of stimuli is also standard in psychoacoustical research (Alexander & Lutfi, 2008; Leibold, Tab, Khaddam, & Jesteadt, 2007). The extent, duration, and rate of these sweeps are in the general range of those observed for frequency glides in speech (Kewley-Port, 1982) as well as the range used in most psychophysical and neuroimaging studies of directional FM sweeps (Hsieh & Saberi, 2010; Brechmann & Scheich, 2005; Poeppel et al., 2004; Saberi & Hafter, 1995).

Two boundary tones were also added to the five-tone FM complex to eliminate cues related to a shift in the centroid of spectral energy between the start and end points of the complex (Figure 1). These two boundary components were pure tones, one of which started at the lowest frequency of the lowest FM component (i.e., 500 Hz) and remained at this frequency for the duration of the stimulus and the other boundary tone started at the highest frequency of the highest FM component (1578 Hz) and remained at that frequency for the duration of the stimulus. To eliminate energy cues associated with summing of boundary tones with the FM component nearest to that tone at the point where they merged, the amplitudes of the lowest and highest FM components of the complex as well as the boundary tones were attenuated by 6 dB (half amplitude) using a logarithmic ramp as the frequency of the FM component approached that of the boundary tone. In addition, to reduce the likelihood of absolute frequency cues affecting neural discrimination, the base frequency (500 Hz; and hence proportionately all other components of the complex) was perturbed on each presentation by up to  $1/3$  of an octave using a randomization routine that allowed the base frequency to start at any of 10 frequencies spaced  $1/30$  of an octave between 500 and 630 Hz (i.e., the start frequency of the next higher-FM component). All stimuli had a 10-msec linear rise decay ramp to reduce onset/offset spectral splatter. Stimulus level was calibrated to 75 dB SPL using a 6-cc coupler, 0.5-in. microphone (Brüel & Kjær, Model 4189), and a precision sound analyzer (Brüel & Kjær, Model 2260). All sounds were presented diotically to both ears. Stimulus delivery and timing were controlled using Cogent software ([www.vislab.ucl.ac.uk/cogent\\_2000.php](http://www.vislab.ucl.ac.uk/cogent_2000.php)) implemented in Matlab (Mathworks, Inc., Natick, MA).



**Figure 1.** A schematic of the FM stimuli. Each complex contained five logarithmic FM sweeps with  $1/3$  octave spacing. The lowest starting frequency was perturbed on each presentation while maintaining the spacing between components. Two boundary tones were also added to the complex to eliminate cues related to a shift in the centroid of spectral energy between the start and end points of the complex (see text for details).

## Experimental Design

There were five experimental conditions: 2 sweep directions  $\times$  2 sweep durations, and a “silence” condition during

which no stimuli were presented. Each condition consisted of a 15-sec block, resulting in a 75-sec-long segment during which all five conditions were presented. This procedure was repeated five times within a run, with a different order of randomization of the five conditions for each repeated 75-sec segment, resulting in a run duration of 375 sec (6 min and 15 sec). Each participant completed five runs with a randomized order of conditions, blocks, and runs across subjects.

Except for the “silence” block, each 15-sec stimulus block was itself subdivided into five subblocks. Each subblock consisted of 2 sec of acoustic stimulation followed by 1 sec of data acquisition (scanning). All auditory stimuli were therefore presented in the silent interval between scanning periods in a sparse-imaging design. For sweep duration of 100 msec, 20 stimuli were concatenated within a 2-sec subblock. For a sweep duration of 400 msec, five stimuli were concatenated within the 2-sec subblock. Within a 15-sec block, the sweep direction and rate were held constant, but the base frequency was perturbed within and across subblocks as described above. The order of randomization of base frequency was different for each subject. Our tests outside the magnet before data collection showed that subjects could not behaviorally discriminate the direction of sweeps within the complex. This was done by design to reduce top-down influences on cortical pattern classification of stimulus features. Nonetheless, participants were asked to listen to the stimuli and try to detect if the direction of the sweep within a block was up or down, but no overt responses were required. This was done only to increase attentiveness to stimuli. Timing of all stimuli was checked for accuracy using a dual-channel digital storage oscilloscope (Tektronix, Model TDS210).

### Scanning Parameters

MR images were obtained in a Siemens 3T (TIM Trio, Siemens, Erlangen, Germany) fitted with a 12-channel radio frequency receiver head coil, at the National Yang-Ming University, Taiwan. To maximize spatial resolution, we collected a total of 126 EPI volumes per each of six sessions in a restricted slice set surrounding the sylvian fissure, parallel to the supratemporal plane, using a T2\*-weighted gradient-echo EPI sequence (matrix =  $108 \times 108$  mm, time repetition [TR] = 3000 msec, acquisition time = 1000 msec, time echo = 34 msec, size =  $2.0 \times 2.0 \times 1.9$  mm, flip angle = 90, number of slices = 11). For a subset of subjects ( $n = 16$ ), we also collected whole-brain EPI volumes in register with the restricted slice set to allow for spatial normalization of functional images; imaging parameters were identical, with the exception of number of slices = 80 and TR = 8860 msec. After the functional scans, a high-resolution anatomical image (T1-weighted) was acquired using a standard MPAGE sequence (matrix =  $224 \times 256$  mm, TR =

2.53 sec, time echo = 3.49 msec, inversion time = 1.1 sec, flip angle = 7, size =  $1 \times 1 \times 1$  mm).

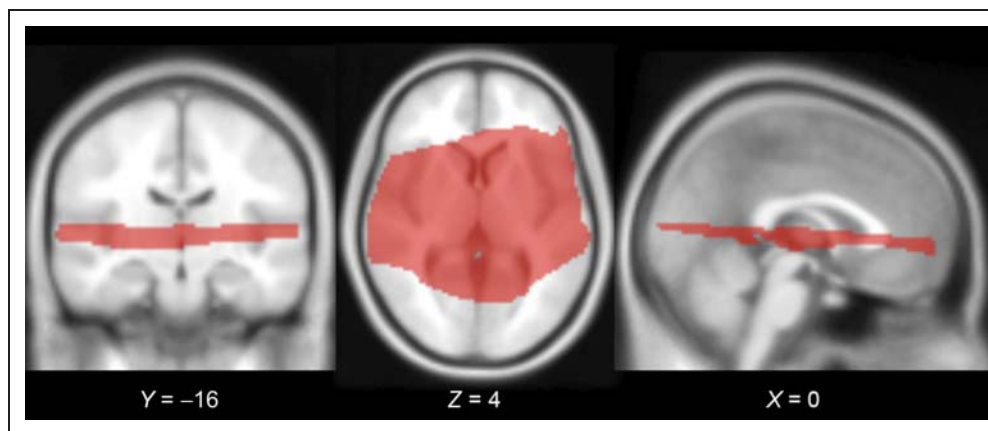
### Data Analysis

Data were analyzed via SPM5 (Wellcome Trust Centre for Neuroimaging, University College, London, UK; implemented in Matlab; Friston et al., 1995). Images were processed in two streams based on intended analysis goals: (1) Individual subject data were modeled for use in identifying ROIs for input to pattern classification, and (2) group data were modeled for a standard analysis using the general linear model (GLM). For both streams, the time series for voxels within each slice was temporally realigned to the middle slice to correct for differences in acquisition time of different slices in a volume. All volumes were realigned spatially to the mean image representing all volumes in the session to correct for any subject motion during the session. Individual subject data were not preprocessed any further. For data intended for group analysis, resulting volumes were spatially normalized to a standard EPI template based on the Montreal Neurological Institute (MNI) reference brain and resampled to 2-mm isotropic voxels. The normalized images were smoothed with a 5-mm FWHM Gaussian kernel and detrended using the linear model of the global signal (Macey, Macey, Kumar, & Harper, 2004).

Statistical analysis was performed in two stages of a mixed effects model. In the first stage, neural activity was modeled by a boxcar function with “on” period corresponding to the duration of the stimulus block. The ensuing BOLD response was modeled by convolving this boxcar function with a canonical hemodynamic response function (Friston et al., 1998), and the time series in each voxel was high-pass-filtered at 1/128 Hz to remove low-frequency noise. Separate covariates were modeled for each condition of interest. Also included for each session was a single covariate representing the mean (constant) over scans, as well as six additional regressors output from the motion correction algorithm. Additionally, two nuisance regressors were included to mark outlier volumes: the first marked volumes, which differed in intensity by more than 2.5 standard deviations from the run mean, and the second marked volumes, which had a large number of voxels which differed from the within-volume mean intensity by more than 2.5 standard deviations. Parameter estimates for events of interest were estimated using a GLM, and effects of interest were tested using linear contrasts of the parameter estimates. For individual subject ROI identification, a liberal threshold of  $p < .001$  (uncorrected), with an extent threshold of 20 voxels was used. Note that the goal for this stage was simply to identify areas likely to be involved in task performance (to be used as input to a classification analysis) rather than to draw direct inference, and as such, correction for multiple comparisons was deemed to be of lesser importance than choosing a threshold which



**Figure 2.** Slice orientation for data collection. To maximize spatial resolution in a minimal period (TR = 3000 msec, acquisition time = 1000 msec, voxel size =  $2.0 \times 2.0 \times 1.9$  mm), we collected EPI volumes in a restricted slice set surrounding the sylvian fissure. Group-level statistical inference was restricted to the subset of voxels for which all subjects' collected slices overlapped when transformed to standard space; this subset is shown here on the SPM MNI152 template, with MNI coordinates listed for each slice.



yielded sufficient numbers of voxels for each subject in the ROIs.

In the group analysis, the contrasts were carried forward to a second stage in which subjects were treated as a random effect. Group-level statistical inference was restricted to the subset of voxels for which all subjects' collected slices overlapped when transformed to standard space, as shown in Figure 2. Contrasts of interest were thresholded at  $p < .05$ , using a false discovery rate (FDR) correction for multiple comparisons (Genovese, Lazar, & Nichols, 2002), as well as a cluster extent threshold of 20 voxels.

## Multivariate Pattern Analysis

### ROI Identification

Data were modeled at the individual subject level to identify subject-specific ROIs for pattern classification analyses. To ensure independence between the ROI selection process and the subsequent testing of contrasts between experimental conditions, ROIs were selected based on the contrast of all experimental conditions versus rest, and subsequent MVPA analyses were conducted only on contrasts between pairs of experimental conditions, which were fully orthogonal to the ROI selection process. Each contrast was also balanced due to the equal number of trials in each experimental condition.

ROI identification was achieved both functionally and anatomically. For each subject, we first identified the cluster of greatest spatial extent, which for almost all subjects encompassed the area in and around Heschl's gyrus (HG) in each hemisphere. We then partitioned the clusters of activity into three ROIs: HG, the anterior STG (aSTG), and the posterior STG (pSTG), the latter two being defined relative to HG as identified on each subject's own anatomy. To count as a cluster in a given sector, there had to be a local maximum within that sector; that is, if a significant cluster extended from HG to aSTG but had

its peak in aSTG, it was counted as a cluster in aSTG, not HG. In the case of multiple local maxima, activations were considered to be in the region containing the majority of voxels within the cluster. Using this method, a majority of subjects had peaks in left and right HG (Lt = 21/21; Rt = 19/21), with fewer having peaks in left and right pSTG (Lt = 9/21; Rt = 11/21), and the least in the left and right aSTG (Lt = 7/21; Rt = 5/21). Across all clusters, the average ROI size was 464 voxels ( $SD = 287$ ) in the left hemisphere and 329 voxels ( $SD = 233$ ) in the right hemisphere (see Table 1 for detailed summaries for each ROI).

### Pattern Classification

ROI-based MVPA (Okada et al., 2010) was implemented in the six ROIs identified in individual subjects to explore the patterns of activation to FM stimuli that vary primarily in terms of sweep direction, but also sweep duration/rate,<sup>1</sup> with the goal of detecting sensitivities in the neural signal that go undetected in a traditional GLM-based analysis. MVPA was achieved using a custom-developed Matlab toolbox based on LIBSVM, a publicly available support vector machine (SVM) library (Chang & Lin, 2011). The logic behind this approach is that if a classifier is able to successfully classify one condition from another based on the pattern of responses in an ROI, then the ROI must contain information that distinguishes the two conditions. In each ROI, four different pairwise classifications were performed: (1) up versus down sweep for all stimuli, (2) up versus down for fast sweeps (100 msec), (3) up versus down for slow sweeps (400 msec), and (4) fast versus slow sweeps (100 vs. 400 msec).

Preprocessing procedures of the signals before applying SVM included standardization and temporal averaging. First, we standardized the motion-corrected and spatially aligned fMRI time series in each recording session (run) by calculating voxel-based  $z$  scores. Second, the standardized data were averaged across the volumes within each block. In addition, to ensure that overall amplitude

**Table 1.** Results of Statistical Classification Analyses for the 21 Subjects in Each ROI and for Each Pairwise Classification Contrast

	<i>Left Hemisphere</i>			<i>Right Hemisphere</i>		
	<i>pSTG</i>	<i>HG</i>	<i>aSTG</i>	<i>pSTG</i>	<i>HG</i>	<i>aSTG</i>
<i>n</i>	9	21	7	11	19	5
Number of voxels ( <i>SD</i> )	316.67 (194.92)	610.24 (393.90)	207.14 (85.14)	261.82 (155.70)	424.21 (334.65)	116.00 (14.22)
Up vs. Down						
Accuracy ( <i>SD</i> )	0.489 (0.110)	0.505 (0.093)	<b>0.541 (0.084)</b>	0.522 (0.091)	<b>0.529 (0.103)</b>	0.510 (0.095)
<i>t</i>	−0.428	0.449	<b>3.023</b>	1.292	<b>2.247</b>	0.328
<i>p</i>	.6601	.3293	<b>.0116</b>	.1127	<b>.0187</b>	.3797
Fast (Up vs. Down)						
Accuracy ( <i>SD</i> )	0.542 (0.134)	0.511 (0.121)	0.489 (0.111)	0.515 (0.116)	0.505 (0.118)	0.520 (0.132)
<i>t</i>	1.701	0.865	−0.431	0.786	0.363	0.520
<i>p</i>	.0637	.1986	.6594	.2251	.3604	.3153
Slow (Up vs. Down)						
Accuracy ( <i>SD</i> )	0.489 (0.110)	0.505 (0.093)	<b>0.527 (0.086)</b>	0.512 (0.090)	<b>0.522 (0.104)</b>	0.510 (0.095)
<i>t</i>	−0.428	0.449	<b>2.554</b>	0.781	<b>1.754</b>	0.328
<i>p</i>	.6601	.3293	<b>.0216</b>	.2263	<b>.0482</b>	.3797
Fast vs. Slow						
Accuracy ( <i>SD</i> )	<b>0.606 (0.131)</b>	<b>0.672 (0.136)</b>	<b>0.626 (0.114)</b>	<b>0.585 (0.115)</b>	<b>0.603 (0.136)</b>	0.524 (0.106)
<i>t</i>	<b>3.178</b>	<b>7.142</b>	<b>3.921</b>	<b>3.476</b>	<b>4.130</b>	0.739
<i>p</i>	<b>.0065</b>	<b>&lt;.0001</b>	<b>.0039</b>	<b>.0030</b>	<b>.0003</b>	.2506

Significant values are shown in **bold**.

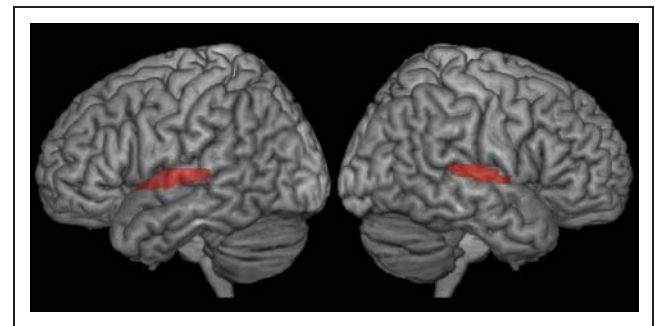
differences between the conditions were not contributing to significant classification, the mean activation level across the voxels within each block was removed before classification. We then performed classification on the pre-processed data set using a leave-one-out cross-validation approach (Vapnik, 1995). In each leave-one-out iteration, we used data from all but one of the five sessions to train the SVM classifier and then used the classifier to test the data from the remaining session. The SVM-estimated condition labels for the testing data set were then compared with the real labels to compute a classification accuracy score. Classification accuracy for each contrast and each subject was derived from averaging the accuracy scores across the five leave-one-out sessions. To assess the significance of classification on the group level for each contrast, we performed a *t* test (one-tailed), comparing the mean accuracy score across subjects to the chance accuracy score (0.5 in the pairwise case; Hickok, Okada, & Serences, 2009).

## RESULTS

### Standard Analysis

Figure 3 shows group results of a standard analysis in which activity patterns in response to all auditory stimuli were

contrasted to rest (scanner noise). Robust auditory activity was observed in the supratemporal plane and regions of the STG ( $p < .05$  [FDR-corrected], extent threshold of 20 voxels). However, when cortical activity in response to up FM sweeps was contrasted to that of down sweeps or when activity associated with the two FM rates were contracted to each other, no significant effects were observed ( $p(\text{FDR}) < .05$ ,  $\text{ET} = 5$ ).



**Figure 3.** Group effects for all auditory conditions versus rest determined from a standard GLM contrast. Robust activations were found in both hemispheres throughout the STG and the supratemporal plane, including HG.

**Figure 4.** Using standard contrast analysis at a threshold of  $p\text{-FDR} < .05$  ( $ET = 20$ ), no clusters were found that were sensitive to either sweep direction or rate. However at a liberal threshold of  $p < .001$  (uncorrected),  $ET = 5$ , several small clusters were more responsive to 100-msec sweeps than 400-msec sweeps. No clusters were found for the reverse contrast (400 msec > 100 msec), nor were any clusters found that were responsive to sweep direction.

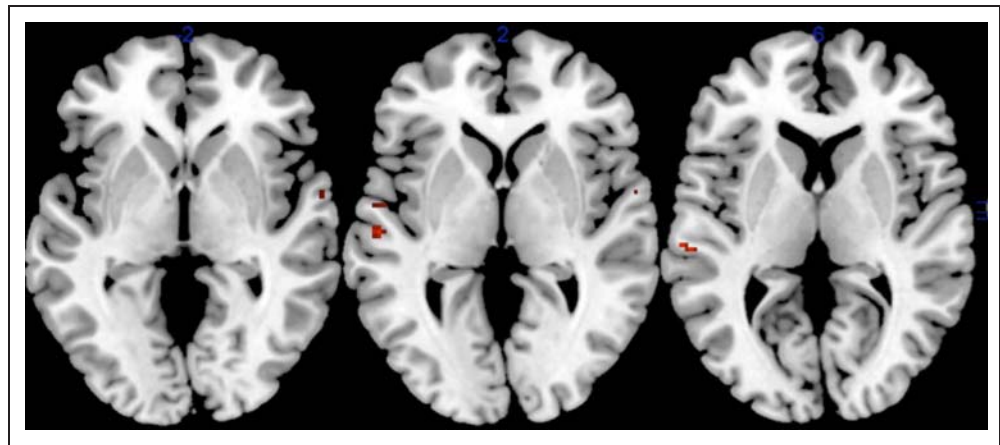


Figure 4 shows that at a liberal threshold of  $p(\text{uncorrected}) < .001$ ,  $ET = 5$ , there were several clusters more responsive to the faster/shorter FM sweep (3.3 octaves/sec) than the slower/longer sweep (0.83 octave/sec), but no areas showed the reverse pattern (slow > fast), and none were sensitive to sweep direction even at this relaxed threshold. Results of our standard contrast analysis therefore failed to identify regions selective to FM coding (at least for sweep direction) and only demonstrated marginal effects of FM duration/rate.

## MVPA

For MVPA, ROIs were identified in individual subject data. We used functionally defined ROIs as input for the pattern classification based on areas active at  $p < .001$  (uncorrected) and  $ET$  of 20 voxels for the average of “All Conditions” minus “Rest.” Clusters were categorized by peak voxel location and general spatial extent in the supratemporal plane relative to HG, labeled as aSTG, HG, or pSTG, in both left and right hemispheres, yielding six total ROIs. Representative slices illustrating ROIs for three subjects are plotted on each subject’s native anatomical image in Figure 5. These regions then were used as “candidate” regions for identification of FM-selective networks that we could assess using a more fine-grained MVPA. The pattern of activity in each subject’s ROI was then assessed in its ability to classify the four different pairs of stimulus conditions noted earlier. Table 1 provides results of statistical classification analyses for the 21 subjects in each ROI (left and right HG, aSTG, pSTG) and for each pairwise classification contrast. Significant results are shown in bold font. MVPA successfully classified FM sweep direction in the right primary auditory cortex and the left aSTG, although the strongest effects were found in the right HG. Further analysis of the sweep direction effect showed that this effect is primarily caused by the longer-duration (slower rate) sweep in both the left aSTG and right HG as classifications in these regions mirror those associated with the full-set contrast between up and down

sweeps, although a larger number of subjects showed significant classification effects for the full-set contrast of up versus down sweeps.<sup>2</sup> Although no significant effect of sweep direction was observed for the faster (100 msec) sweep, the classification of voxel activation patterns in the left pSTG for this stimulus did approach significance ( $t = 1.70$ , accuracy = 0.542,  $p = .06$ ) consistent with two prior studies that have implicated pSTG in FM processing (Brechmann & Scheich, 2005; Johnsrude et al., 1997). The classifier also successfully distinguished between slow and fast FM rates bilaterally in the primary auditory cortex and in the left anterior and posterior regions of STG as well as pSTG in the right hemisphere. For illustrative purposes, classification results are also overlaid on a schematic brain image in Figure 6.

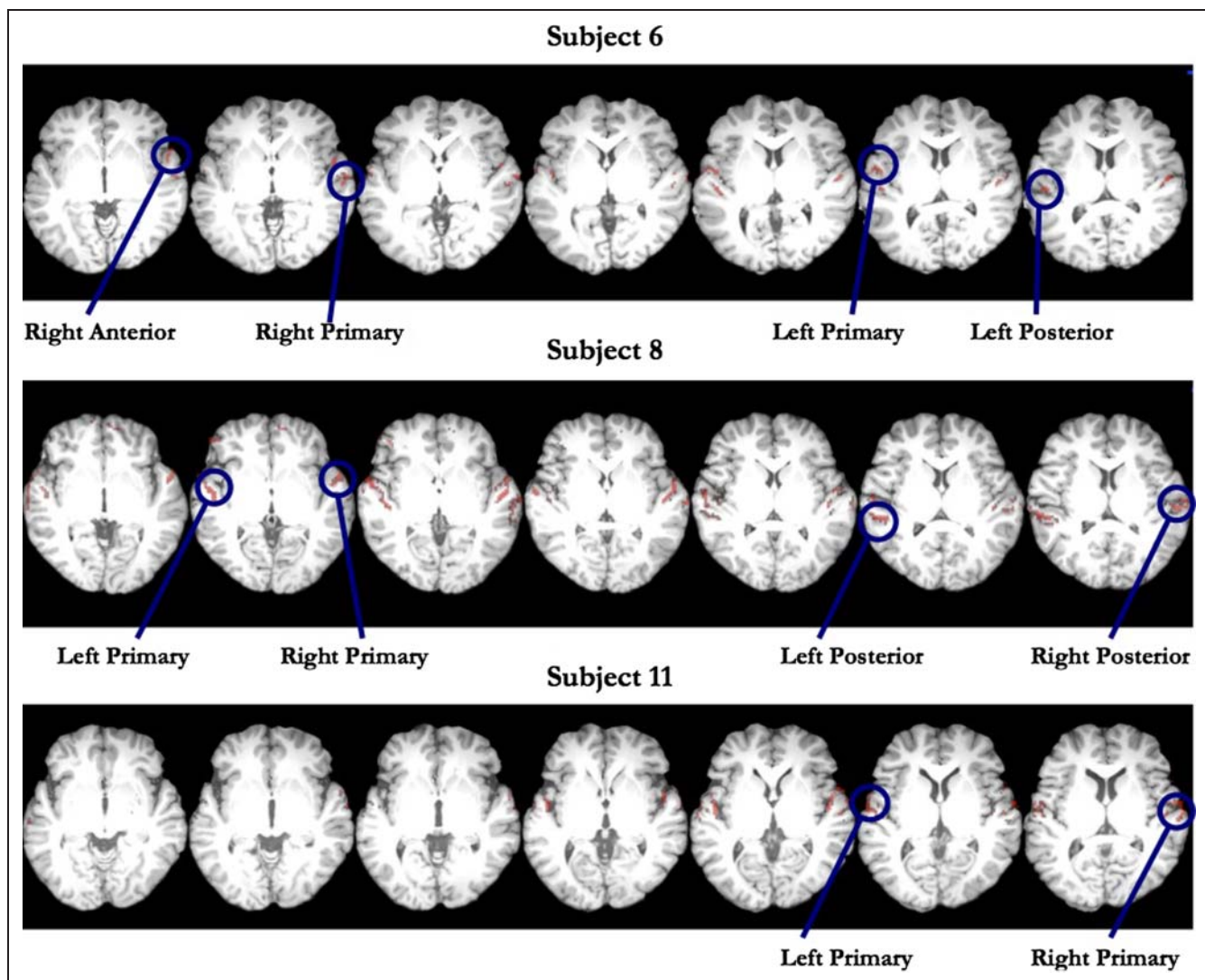
## DISCUSSION

The current study used multivariate pattern classification analysis to reveal existence of FM-sensitive direction-selective networks in human primary and secondary auditory cortex. Standard contrast analysis failed to identify direction-selective regions, likely because the assembly of neurons that codes for one direction of FM is interspersed within the area that codes for the other direction. This finding also suggests that the number of cells that process one sweep direction is approximately the same as that for the other within ROIs, a symmetry feature that is not necessarily true for animal species whose auditory cortices have either significantly larger numbers of down-sweep neurons (Andoni et al., 2007; Voytenko & Galazyuk, 2007; Fuzessery, Richardson, & Coburn, 2006) or which have asymmetric differences in distribution of up- and down-sweep cells in different ROIs (Godey et al., 2005; Tian & Rauschecker, 2004).

Animal neurophysiological studies suggest that the first site of FM coding is early in the auditory brainstem in the inferior colliculus (Williams & Fuzessery, 2010) and possibly earlier in the system (Gittelman et al., 2009). No FM

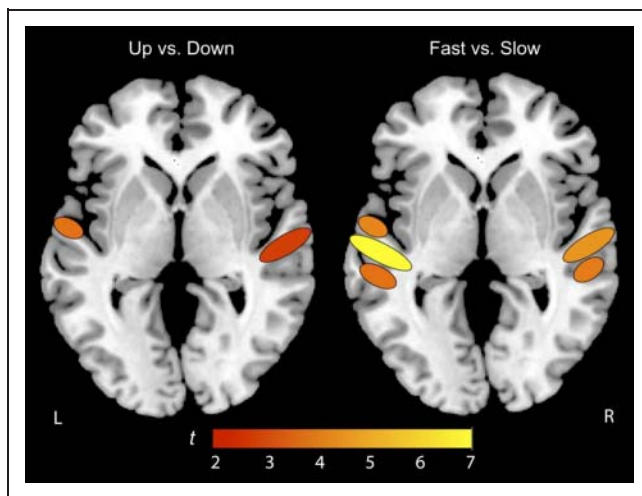
selectivity exists in the auditory nerve. Our study focused on higher brain centers that may contain such putative networks of FM-selective neurons and is the first to identify brain regions selectively responsive to a within-stimulus class feature of FM sounds (i.e., sweep direction). These findings are different than those based on standard GLM subtraction methods which contrast broader task categories (e.g., judgment of sweep direction contrasted to phonemic identification). Differences in neural activity associated with the latter contrast can originate from differences in decision processes from differences in classes of stimuli employed (e.g., phonemes, tones, narrowband noises, FM or AM sounds) or from processing of sweep direction. It is probable that a combination of these factors have contributed to the reported patterns of cortical activity. Our analysis using GLM subtraction methods did not

identify any regions selective for sweep direction, suggesting that findings from prior neuroimaging studies of FM selectivity are not likely to have originated from processing of FM cues per se but rather from contrasts between broader stimulus categories. Our findings, additionally, point to the critical importance of using advanced statistical classification methods sensitive to subtle differences in voxel activation patterns for the study of stimulus dimensions not associated with a simple monotonic change in the location of peak neural activity (e.g., tonotopic organization; Woods et al., 2009). More complex stimulus categories, which comprise most sounds of interest, will likely lead to intricate changes in patterns of neural activity within a region (Abrams et al., 2011; Stecker, Harrington, & Middlebrooks, 2005; Stecker & Middlebrooks, 2003) that would remain undetected by standard contrast methods.



**Figure 5.** Example clusters of interest from three subjects used as inputs for pattern classification. Clusters were defined based on all conditions versus rest, at a liberal threshold of  $p < .001$  (uncorrected) with an extent threshold of 20 voxels. For each subject, we first identified the cluster of greatest spatial extent, generally in and around HG in each hemisphere. We then subdivided the STG into HG, aSTG, and pSTG regions and identified peaks of activity that fell within these anatomically defined sectors in both the left and right hemispheres. A majority of subjects had peaks in the left and right HG, with fewer having peaks in pSTG, and with the least number of peaks occurring in aSTG.





**Figure 6.** Schematic of pattern classification results. MVPA successfully classified FM sweep direction in the right primary auditory cortex and the left aSTG. The classifier also successfully distinguished between slow and fast FM rates bilaterally in the primary auditory cortex and in the left anterior and posterior regions of STG as well as pSTG in the right hemisphere.

In our study, we identified two brain regions, the right primary auditory cortex and anterior regions of the left STG in the secondary auditory cortex that are sensitive to direction of an FM sweep. Prior human neuroimaging studies provide a mixed picture of brain areas responsive to FM tasks. Nearly all regions of the auditory cortex (bilaterally) have in one or another study been implicated as FM specialist. These include the right auditory cortex in general (König, Sieluzicky, Simserides, Heil, & Scheich, 2008; Behne et al., 2005; Schlosser et al., 1998), the right posterior auditory cortex (Brechmann & Scheich, 2005; Poeppel et al., 2004), the left primary auditory cortex (Hall et al., 2002; Belin et al., 1998), the left posterior auditory cortex (Brechmann & Scheich, 2005; Johnsrude et al., 1997), the left aSTG (Thivard et al., 2000), as well as bilaterally in primary (Belin, 1998) and the secondary auditory cortex posterior to HG (Thivard et al., 2000), posterior supratemporal plane bilaterally (STP, Hall et al., 2002), anterior STP and STS bilaterally (Hall et al., 2002). Regions that are not primarily auditory have also been shown to be responsive to FM tasks when contrasted to other auditory tasks. These areas include the right cerebellum, left fusiform gyrus, left OFC (Johnsrude et al., 1997), right dorsolateral pFC, left superior parietal lobule, and the supramarginal gyrus (Poeppel et al., 2004).

We consider this varied picture to be an inaccurate representation of regions containing FM-specialized networks and more descriptive of top-down cognitive influences in contrasting distinct decision tasks (Brechmann & Scheich, 2005). It is possible that some of these prior studies have perhaps tapped into existing FM networks, although they cannot conclusively or directly demonstrate their existence without appropriate controls for task effects. It would therefore be useful to consider how our

findings compare to the varied findings reported by these studies. We have found that our results are most consistent with those of Thivard et al. (2000), who identified the left aSTG and with those studies that have identified the right primary auditory regions as selective for FM (König et al., 2008; Behne et al., 2005; Schlosser et al., 1998). The left aSTG has been argued to be involved in on-line syntactic processing, speech perception, and sentence comprehension (Matsumoto et al., 2011; Friederici, Meyer, & von Cramon, 2000; Meyer, Friederici, & von Cramon, 2000) including prosodic aspects (Humphries, Love, Swinney, & Hickok, 2005), whereas the right primary auditory cortex is speculated to be involved in spectro-temporal processing of sustained sounds and temporal segmentation of auditory events (Izumi et al., 2011; Poeppel, 2003; Zatorre & Belin, 2001). Whether these two regions, the right aSTG and left primary areas work in concert to integrate FM information into networks associated with speech processing and comprehension is unclear and requires studies that explore the inter-hemispheric time course of activation patterns in these regions using stimuli that control jointly for acoustic features and speech content. Interhemispheric interactions of this type are plausible. Behne et al. (2005), for example, have shown that presentation of contralateral white noise significantly increases activity in the right auditory cortex in response to ipsilateral FM tones, but contralateral noise has no effect on the activation of the left auditory cortex by FM.

The current study was a step toward understanding auditory cortical response patterns to complex FM tones and is the first to directly demonstrate existence of brain networks specifically responsive to an FM stimulus feature. Our findings contribute to the body of information on how biological systems code FM-based signals that are prevalent in a broad category of complex modulated sounds that include but are not limited to speech, music, animal vocalizations, and other natural sounds. The current findings merit additional future investigations of auditory cortex function in response to sounds comprising complex sweeps of simultaneously differing rates and directions as would occur naturally in ecologically relevant sounds.

### Acknowledgments

We thank Tugan Muftuler for helpful discussions and Haleh Farahbod for assistance in manuscript preparation. This work was supported by grants from the National Science Council, Taiwan (NSC 98-2410-H-008-081) to I. H. and the National Institutes of Health (R01DC009659) to G. H. and K. S.

Reprint requests should be sent to I-Hui Hsieh, Institute of Cognitive Neuroscience, National Central University, Jhongli City, Taiwan, or via e-mail: ihsieh@ncu.edu.tw.

### Notes

1. The purpose of the current study was to identify FM networks selective for sweep direction, and to this end, two stimulus

durations were selected to optimize the likelihood of identifying such networks. FM duration, rate, and bandwidth (peak frequency excursion) are necessarily linked in FM sounds in that one only has two degrees of freedom to change values of three parameters (e.g., keeping peak frequency deviation or bandwidth constant while varying duration results in a change in sweep rate). Thus, while we decided here to additionally conduct analyses for sweep durations/rates, we should clarify that duration and rate effects cannot be dissociated without a parametric study of rate, duration, and bandwidth, which is outside the scope and aims of the current study.

2. Some accuracy values, and hence the same  $t$  and  $p$  values, are identical. It is likely that for these specific cases, the 400-msec condition carries more information than the 100-msec condition about sweep direction, and hence plays the dominant role in training the classifier for discriminating the overall sweep direction effect.

## REFERENCES

- Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, *21*, 1507–1518.
- Alexander, J. M., & Lutfi, R. A. (2008). Sample discrimination of frequency by hearing-impaired and normal-hearing listeners. *Journal of the Acoustical Society of America*, *123*, 241–253.
- Andoni, S., Li, N., & Pollak, G. D. (2007). Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. *Journal of Neuroscience*, *27*, 4882–4893.
- Behne, N., Scheich, H., & Brechmann, A. (2005). Contralateral white noise selectively changes right human auditory cortex activity caused by a FM-direction task. *Journal of Neurophysiology*, *93*, 414–423.
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., et al. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, *10*, 536–540.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Brechmann, A., & Scheich, H. (2005). Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cerebral Cortex*, *15*, 578–587.
- Carlyon, R. P. (1994). Detecting mistuning in the presence of synchronous and asynchronous interfering sounds. *Journal of the Acoustical Society of America*, *95*, 2622–2630.
- Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*, 1–27.
- Chen, H. B., & Zeng, F. G. (2004). Frequency modulation detection in cochlear implant subjects. *Journal of the Acoustical Society of America*, *116*, 2269–2277.
- Coscia, E. M., Phillips, D. P., & Fentress, J. C. (1991). Spectral analysis of neonatal wolf vocalizations. *Bioacoustics*, *3*, 275–293.
- d'Alessandro, C., Rosset, S., & Rossi, J. P. (1998). The pitch of short-duration fundamental frequency glissandos. *Journal of the Acoustical Society of America*, *104*, 2339–2348.
- Dankiewicz, L. A., Helweg, D. A., Moore, P. W., & Zafran, J. M. (2002). Discrimination of amplitude-modulated synthetic echo trains by an echolocating bottlenose dolphin. *Journal of the Acoustical Society of America*, *112*, 1702–1708.
- Dear, S. P., Simmons, J. A., & Fritz, J. (1993). A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat. *Nature*, *364*, 620–623.
- Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, *27*, 769–773.
- Divenyi, P. L. (2009). Perception of complete and incomplete formant transitions in vowels. *Journal of the Acoustical Society of America*, *126*, 1427–1439.
- Divenyi, P. L., & Robinson, A. J. (1989). Non-linguistic auditory capabilities in aphasia. *Brain and Language*, *37*, 290–326.
- Fant, G. C. M. (1970). *Acoustic theory of speech production*. The Hague, Netherlands: Mouton.
- Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, *74*, 289–300.
- Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *3*, 165–189.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage*, *7*, 30–40.
- Fuzessery, Z. M., & Hall, J. C. (1996). Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. *Journal of Neurophysiology*, *76*, 1059–1073.
- Fuzessery, Z. M., Richardson, M. D., & Coburn, M. S. (2006). Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the inferior colliculus of the pallid bat. *Journal of Neurophysiology*, *96*, 1320–1336.
- Genovese, R. C., Lazar, N. A., & Nichols, T. E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, *15*, 870–878.
- Gittelman, J. X., Li, N., & Pollak, G. D. (2009). Mechanisms underlying directional selectivity for frequency-modulated sweeps in the inferior colliculus revealed by in vivo whole-cell recordings. *Journal of Neuroscience*, *29*, 13030–13041.
- Godey, B., Atencio, C. A., Bonham, B. H., Schreiner, C. E., & Cheung, S. W. (2005). Functional organization of squirrel monkey primary auditory cortex: Responses to frequency-modulation sweeps. *Journal of Neurophysiology*, *94*, 1299–1311.
- Gordon, M., & O'Neill, W. E. (1998). Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. *Hearing Research*, *122*, 97–108.
- Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *12*, 140–149.
- Hickok, G., Okada, K., & Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology*, *101*, 2725–2732.
- Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*. Cambridge, England: Cambridge University Press.
- Hsieh, I., & Saberi, K. (2009). Detection of spatial cues in linear and logarithmic frequency-modulated sweeps. *Attention, Perception & Psychophysics*, *71*, 1876–1889.

- Hsieh, I., & Saberi, K. (2010). Detection of sinusoidal amplitude modulation in logarithmic frequency sweeps across wide regions of the spectrum. *Hearing Research*, 262, 9–18.
- Huber, F., & Thorson, J. (1985). Cricket auditory communication. *Scientific American*, 253, 60–68.
- Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping*, 26, 128–138.
- Izumi, S., Itoh, K., Matsuzawa, H., Takahashi, S., Kwee, I. L., & Nakada, T. (2011). Functional asymmetry in primary auditory cortex for processing musical sounds: Temporal pattern analysis of fMRI time series. *NeuroReport*, 22, 470–474.
- Johnsrude, I. S., Zatorre, R. J., Milner, B. A., & Evans, A. C. (1997). Left-hemisphere specialization for the processing of acoustic transients. *NeuroReport*, 8, 1761–1765.
- Kajikawa, Y., De la Mothe, L. A., Blumell, S., Sterbing-D'Angelo, S. J., D'Angelo, W. R., Camalier, C., et al. (2008). Coding of FM sweep trains and twitter calls in area CM of marmoset auditory cortex. *Hearing Research*, 239, 107–125.
- Kewley-Port, D. (1982). Measurement of formant transitions in naturally produced stop consonant-vowel syllables. *Journal of the Acoustical Society of America*, 72, 379–389.
- Klump, G. M., & Langemann, U. (1992). The detection of frequency and amplitude modulation in the European starling *Sturnus vulgaris* psychoacoustics and neurophysiology. *Advances in the Biosciences*, 83, 353–359.
- Koch, U., & Grothe, B. (1998). GABAergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat. *Journal of Neurophysiology*, 80, 71–82.
- König, R., Sieluzycski, C., Simserides, C., Heil, P., & Scheich, H. (2008). Effects of the task of categorizing FM direction on auditory evoked magnetic fields in the human auditory cortex. *Brain Research*, 1220, 102–117.
- Leibold, L. J., Tab, H. Y., Khaddam, S., & Jesteadt, W. (2007). Contributions of individual components to the overall loudness of a multitone complex. *Journal of the Acoustical Society of America*, 121, 2822–2831.
- Lieberman, A. M., Delattre, P. C., Gerstman, L. J., & Cooper, F. S. (1956). Tempo of frequency change as cue for distinguishing classes of speech sounds. *Experimental Psychology*, 52, 127–137.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368.
- Luo, H., Wang, Y., Poeppel, D., & Simon, J. Z. (2007). Concurrent encoding of frequency and amplitude modulation in human auditory cortex: Encoding transition. *Journal of Neurophysiology*, 98, 3473–3485.
- Macey, P. M., Macey, K. E., Kumar, R., & Harper, R. M. (2004). A method for removal of global effects from fMRI time series. *Neuroimage*, 22, 360–366.
- Matsumoto, R., Imamura, H., Inouchi, M., Nakagawa, T., Yokoyama, Y., Matsushashi, M., et al. (2011). Left anterior temporal cortex actively engages in speech perception: A direct cortical stimulation study. *Neuropsychologia*, 49, 1350–1354.
- Meyer, M., Friederici, A. D., & von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: Event-related fMRI reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research*, 9, 19–33.
- Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I., Saberi, K., et al. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex*, 20, 2486–2495.
- Pickett, J. M. (1980). *The sounds of speech communication*. Baltimore: University Park Press.
- Pickles, J. O. (2008). *An introduction to the physiology of hearing* (3rd ed.). Bradford, UK: Emerald Group Publishing.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as “asymmetric sampling in time.” *Speech Communication*, 41, 245–255.
- Poeppel, D., Guillemin, A., Thompson, J., Fritz, J., Bavelier, D., & Braun, A. R. (2004). Auditory lexical decision categorical perception and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia*, 42, 183–200.
- Razak, K., & Fuzessery, Z. (2006). Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex of the pallid bat. *Journal of Neurophysiology*, 96, 1303–1319.
- Razak, K. A., & Fuzessery, Z. M. (2010). Experience-dependent development of vocalization selectivity in the auditory cortex. *Journal of the Acoustical Society of America*, 128, 1446–1451.
- Ryan, M. J., & Wilczynskin, W. (1988). Convolution of sender and receiver: Effect on local mate preference in cricket frogs. *Science*, 240, 1786–1788.
- Saberi, K., & Hafer, E. R. (1995). A common neural code for frequency- and amplitude-modulated sounds. *Nature*, 374, 537–539.
- Sabourin, P., Gottlieb, H., & Pollack, G. S. (2008). Carrier-dependent temporal processing in an auditory interneuron. *Journal of the Acoustical Society of America*, 123, 2910–2917.
- Scheich, H., Baumgart, F., Gaschler-Markefski, B., Tegeler, C., Tempelmann, C., Heinze, H. J., et al. (1998). Functional magnetic resonance imaging of a human auditory cortex area involved in foreground-background decomposition. *European Journal of Neuroscience*, 10, 803–809.
- Schlosser, M., Aoyagi, N., Fulbright, R., Gore, J., & McCarthy, G. (1998). Functional MRI studies of auditory comprehension. *Human Brain Mapping*, 6, 1–13.
- Stagray, J. R., Downs, D., & Sommers, R. K. (1992). Contributions of the fundamental resolved harmonics and unresolved harmonics in tone-phoneme identification. *Journal of Speech and Hearing Research*, 35, 1406–1409.
- Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location coding by opponent neural populations in the auditory cortex. *PLOS Biology*, 3, 520–528.
- Stecker, G. C., & Middlebrooks, J. C. (2003). Distributed coding of sound locations in the auditory cortex. *Biological Cybernetics*, 89, 341–349.
- Subramanian, A., Yairi, E., & Amir, O. (2003). Second formant transitions in fluent speech of persistent and recovered preschool children who stutter. *Journal of Communication Disorders*, 36, 59–75.
- Suga, N. (1968). Analysis of frequency-modulated and complex sounds by single auditory neurones of bats. *Journal of Physiology*, 198, 51–80.
- Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X. Q., Nagarajan, S. S., et al. (1996). Language comprehension in language learning impaired children improved with acoustically modified speech. *Science*, 271, 81–84.
- Thivard, L., Belin, P., Zilbovicius, M., Poline, J. B., & Samson, Y. (2000). A cortical region sensitive to auditory spectral motion. *NeuroReport*, 11, 2969–2972.
- Tian, B., & Rauschecker, J. P. (2004). Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *Journal of Neurophysiology*, 92, 2993–3013.
- Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.

- Voytenko, S. V., & Galazyuk, A. V. (2007). Intracellular recording reveals temporal integration in inferior colliculus neurons of awake bats. *Journal of Neurophysiology*, 97, 1368–1378.
- Wetzel, W., Ohl, F. W., Wagner, T., & Scheich, H. (1998). Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. *Neuroscience Letters*, 252, 115–118.
- Williams, A. J., & Fuzessery, Z. M. (2010). Facilitatory mechanisms shape selectivity for the rate and direction of FM sweeps in the inferior colliculus of the pallid bat. *Journal of Neurophysiology*, 104, 1456–1471.
- Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., et al. (2009). Functional maps of human auditory cortex: Effects of acoustic features and attention. *PLOS One*, 4, e5183.
- Woolley, S. M. N., & Casseday, J. H. (2005). Processing of modulated sounds in the zebra finch auditory midbrain: Responses to noise frequency sweeps and sinusoidal amplitude modulations. *Journal of Neurophysiology*, 94, 1143–1157.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946–953.