Behavioral/Systems/Cognitive

Multiplexed and Robust Representations of Sound Features in Auditory Cortex

Kerry M. M. Walker, Jennifer K. Bizley, Andrew J. King, and Jan W. H. Schnupp
Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3PT, United Kingdom

We can recognize the melody of a familiar song when it is played on different musical instruments. Similarly, an animal must be able to recognize a warning call whether the caller has a high-pitched female or a lower-pitched male voice, and whether they are sitting in a tree to the left or right. This type of perceptual invariance to "nuisance" parameters comes easily to listeners, but it is unknown whether or how such robust representations of sounds are formed at the level of sensory cortex. In this study, we investigate whether neurons in both core and belt areas of ferret auditory cortex can robustly represent the pitch, formant frequencies, or azimuthal location of artificial vowel sounds while the other two attributes vary. We found that the spike rates of the majority of cortical neurons that are driven by artificial vowels carry robust representations of these features, but the most informative temporal response windows differ from neuron to neuron and across five auditory cortical fields. Furthermore, individual neurons can represent multiple features of sounds unambiguously by independently modulating their spike rates within distinct time windows. Such multiplexing may be critical to identifying sounds that vary along more than one perceptual dimension. Finally, we observed that formant information is encoded in cortex earlier than pitch information, and we show that this time course matches ferrets' behavioral reaction time differences on a change detection task.

Introduction

The encoding properties of auditory neurons have typically been examined by expressing their spike rates within a fixed time window as a function of a single sound parameter. Rate–intensity curves, for example, describe firing rates as a function of sound intensity (Phillips and Irvine, 1981). This approach has revealed much about how neuronal responses are modulated by sounds but overlooks the inherent complexities of natural sound statistics and spike firing patterns.

Outside of the laboratory, sounds typically vary over more than one acoustic dimension. Thus, an animal must be able to recognize the calls of predators, prey, and potential mates regardless of their pitch or the direction from which they originate. Neuronal sensitivity to timbre (Langner et al., 1981; Ohl and Scheich, 1997), pitch (Bendor and Wang, 2010; Bizley et al., 2010), and spatial location (Benson et al., 1981; Middlebrooks et al., 1994; Mrsic-Flogel et al., 2005) has been described in auditory cortex. However, it is not known whether or how the activity of cortical neurons can support perceptual invariance to "nuisance" parameters when these features vary simultaneously.

One way to achieve an explicit, invariant representation of multiple sound properties is to represent each feature by a separate group of neurons. However, our recent results suggest that this is not the way auditory cortical response properties are organized. Rather, the vast majority of neurons in ferret auditory cortex are sensitive to combinations of the periodicity, timbre, and location of sounds (Bizley et al., 2009). An alternative solution might be to represent multiple cues within the spiking response of individual neurons by tuning different aspects of the response to different perceptual features, resulting in a "multiplexed" spike code. Indeed, the onset, sustained, and offset responses of cortical neurons can represent different stimulus values for a range of sound features, including pure tone frequency and intensity (Takahashi et al., 2004; Wang et al., 2005), spatial location (Campbell et al., 2010), and vocalization identity (Qin et al., 2008). These findings emphasize the importance of choosing an appropriate response window for determining the sensitivity of a neuron to a given sound feature (Panzeri et al., 2010). They are also consistent with the possibility that single neurons hold multiplexed representations of sound features, but this has yet to be confirmed experimentally.

Here, we investigate whether neurons carry information about the periodicity, spectral timbre (formants), or azimuthal location of artificial vowel sounds in ways that are "robust" to simultaneous variance in the other two attributes. We estimated the information carried about each attribute within discrete time windows of the neural response and show that the neural code for timbre, pitch, and spatial location differs systematically across the auditory cortex. Furthermore, we demonstrate that single neurons can multiplex pitch and timbre information in separate response windows. Robust information about sound timbre was found to peak consistently earlier in the response than pitch information. We show that this neural signature has a behavioral correlate: ferrets respond to timbre changes faster than pitch changes in artificial vowels in a go/no-go task.

Received April 26, 2011; revised July 14, 2011; accepted Aug. 9, 2011.

Author contributions: K.M.M.W., J.K.B., A.J.K., and J.W.H.S. designed research; K.M.M.W. and J.K.B. performed research; K.M.M.W. and J.W.H.S. contributed unpublished reagents/analytic tools; K.M.M.W. analyzed data; K.M.M.W. wrote the paper.

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council (J.W.H.S., J.K.B., A.J.K.), a Royal Society Dorothy Hodgkin Fellowship (J.K.B.), and a Wellcome Trust Principal Research Fellowship (A.J.K.). We are grateful to Israel Nelken for his helpful comments on a draft of this manuscript, and Cesare Magri for providing MATLAB code for entropy calculations at http://www.ibtb.org (Magri et al., 2009).

Correspondence should be addressed to Kerry M. M. Walker, Department of Physiology, Anatomy and Genetics, University of Oxford, Sherrington Building, Parks Road, Oxford OX1 3PT, UK. E-mail: kerry.walker@dpag.ox.ac.uk.

DOI:10.1523/JNEUROSCI.2074-11.2011

Copyright © 2011 the authors 0270-6474/11/3114565-12$15.00/0

Materials and Methods

Animals. A total of 10 adult, female, pigmented ferrets (*Mustela putorius*) were used in this study. All experiments were approved by the local ethical review committee and performed under license from the United Kingdom Home Office in accordance with the Animal (Scientific Procedures) Act (1986).

Extracellular recordings. Details of the surgical procedures performed here on six ferrets can be found in the study by Bizley et al. (2009). Briefly, general anesthesia was induced with a single intramuscular dose of medetomidine (Domitor; 0.022 $\mathrm{mg \cdot kg^{-1} \cdot h^{-1}}$; Pfizer) and ketamine (Ketaset; 5 $\mathrm{mg \cdot kg^{-1} \cdot h^{-1}}$; Fort Dodge Animal Health) and was maintained with a continuous intravenous infusion of medetomidine and ketamine in saline throughout the experiment. Oxygen was supplemented with a ventilator, and vital signs (body temperature, end-tidal CO_2, and the electrocardiogram) were monitored while the temporal muscles were retracted, a head holder was secured to the skull surface, and a craniotomy and durotomy were made over the left auditory cortex.

Spike activity was recorded using silicon array Michigan probes (1–2 MΩ impedances; Neuronexus). Electrode penetrations were made throughout five identified areas of ferret auditory cortex: the primary auditory cortex (A1), the anterior auditory field (AAF), the posterior pseudosylvian and suprasylvian fields (PPF and PSF), and the anterior dorsal field (ADF). The electrode recording sites used were configured as 8×4 (eight active sites on four parallel probes, with a vertical spacing of 150 μm), 4×4 (100–150 μm spacing), 16×2 (spaced at 100 μm intervals), or 16×1 (100–150 μm spacing on a single probe). Photographic records of each electrode penetration were used to reconstruct the recording locations relative to anatomical landmarks (surface blood vessels and sulcal patterns).

Voltage signals recorded were bandpass filtered (500 Hz to 5 kHz), amplified up to 20,000 times, and digitized at 25 kHz. Data acquisition was performed using Tucker-Davis Technologies System 3 multichannel recording systems, together with desktop computers running BrainWare software (Tucker-Davis Technologies) and custom scripts written in MATLAB (MathWorks).

Neuronal spiking responses were isolated from the digitized signal off-line. Spikes from a common neural source were classified either by manually clustering spike shapes according to features such as amplitude, width, and area, or by using an automated *k*-means algorithm to cluster voltage potentials. A unit was only classified as a single neuron if the autocorrelation spike histogram revealed a clear refractory period and the spike shape was stereotyped. All other clusters of spikes were classified as multiunit activity, representing the firing of a small cluster of neurons. Poststimulus time histograms (PSTHs) of spikes from each unit were visually inspected in BrainWare, and all single units ($n = 619$) and multiunit clusters ($n = 464$) that appeared to be driven by auditory stimuli were exported to MATLAB for further analysis.

Stimuli. Sounds were generated using Tucker-Davis Technologies System 3 hardware and MATLAB. Sounds were presented to anesthetized animals in an anechoic room through Panasonic earphone drivers (RPHV297), which were mounted on plastic otoscope speculae inserted into each ear canal. The earphones were closed-field calibrated using a 1/8 inch condenser microphone (Brüel and Kjær type 4138) placed at the end of a model ferret ear canal, and inverse filters were used to ensure that the devices produced flat (less than ±5 dB) outputs across 100–24,000 Hz.

Following each cortical penetration, noise bursts (100 ms duration; 900 ms interstimulus interval; 10–80 dB SPL) were presented to identify acoustically responsive neural activity. Next, pure tones (100 ms duration, with 5 ms cosine onset and offset ramps) were presented, and the best frequency (BF) of each unit was derived from the frequency response area, as described by Bizley et al. (2005). All units with a bandwidth of <2 octaves at 10 dB above threshold were assigned a BF, and those with wider bandwidths were classified as "untuned."

Artificial vowel stimuli were created in MATLAB, using an algorithm adapted from Malcolm Slaney's Auditory Toolbox. A click train was first produced with a repetition rate corresponding to the desired fundamental frequency (F0) and a duration of 150 ms. This signal was then passed through a four-pass-band filter to impart spectral peaks at the desired formant frequencies. The artificial vowels were root-mean-square normalized to ensure that changes in pitch or timbre did not influence the overall sound pressure level. Virtual acoustic space (VAS) techniques were then used to add sound source direction cues (interaural timing differences, interaural level differences, and spectral cues) to the artificial vowel sounds, as described by Mrsic-Flogel et al. (2005).

The sounds were presented in VAS at −45, −15, 15, and 45° azimuth, at 0° elevation. Negative azimuths denote locations to the animal's right, contralateral to the recording hemisphere. F0 values of 200, 336, 565, and 951 Hz were used. The four timbres presented corresponded to the following vowels: /a/ with formant frequencies F1–F4 at 936, 1551, 2815, and 4290 Hz; /ɛ/ with formant frequencies at 730, 2058, 2979, and 4294 Hz; /u/ with formant frequencies at 460, 1105, 2735, and 4115 Hz; and /i/ with formant frequencies at 437, 2761, 3372, and 4352 Hz. The combinations of these four pitches, four timbres, and four sound source directions resulted in a stimulus set of 64 artificial vowels.

Mutual information calculations. Mutual information (MI) calculations have been used to evaluate how reliably the response of a neuron distinguishes between sensory stimuli. They provide an estimate of the reduction in uncertainty about a stimulus feature, *X*, provided by the response, *R*, of a neuron. Therefore, when the MI is larger, one can more accurately guess the identity of the stimulus presented based on the neural response. By convention, MI is measured in bits (Cover and Thomas, 1991), according to the following formula:

$$\mathrm{MI}(X;R) = \sum_{x \in X} \sum_{r \in R} p(x,r) \log_2\left(\frac{p(x,r)}{p(x)\,p(r)}\right), \qquad (1)$$

The artificial vowel sounds used in the present experiment were varied over three perceptual attributes: timbre, pitch, and sound source azimuth. We used MI analysis to examine how well a neuron can encode one of these stimulus dimensions (X), despite variance in the other two dimensions (Y and Z). To calculate MI for stimulus feature X (timbre, for example), responses to presentations of the 64 vowel sounds were classified into the four X groups, collapsing across the 16 Y and Z (e.g., F0 and azimuth) combinations. Therefore, in the case of MI calculations for timbre, the value of x in Equation 1 would indicate only the timbre of the sound, and the size of set X was four. We presented each of the 64 sounds in our stimulus set at least 20 times for each unit, so this grouping of stimuli meant that each MI calculation included at least 320 observations of neural responses (i.e., 320 "trials") for each value of x ($N_x = 320$).

The maximum MI for a stimulus set of size $|X|$ with uniform probability is equal to $\log_2(|X|)$. We presented each of our stimulus classes with equal probability, so an ideal neuron that responds perfectly reliably to different values of sound feature X, and invariantly to features Y and Z, would provide $\log_2(4)$, or 2, bits of information.

The response, r, in Equation 1 was the sum of spikes within a single response time window. The response window was defined by its duration and onset time. MI was calculated for several window durations, including 20, 40, and 320 ms. For each response window duration, MI calculations were repeated across a range of onset times with respect to sound onset. The window onset was shifted across the first 320 ms of the neural response, in steps of one-half of the window size. For a window width of 20 ms, this would give 33 estimates of MI, with onsets ranging from 0 to 320 ms. For a response window of 320 ms, the MI was calculated once, summing spikes occurring on each trial from sound onset ($t = 0$ ms) to 320 ms. For each MI calculation, the set R comprised all unique neural responses, r, observed from the neuron, across all stimulus presentations.

The information content of codes other than spike rates was also examined. For a binary spiking code, trials with at least one spike in the 320 ms following sound onset were coded as "1," while trials in which the unit did not spike were coded as "0." We also wanted to examine the information available in the first-spike latency to millisecond accuracy, although this value was often too variable (i.e., the entropy of first-spike latency was too high) to adequately estimate the probability densities. Therefore, we restricted the variability of this code by binning the latencies into four first-spike latency quartiles for each unit. Here, trials on which there were no spike within the 320 ms response window were

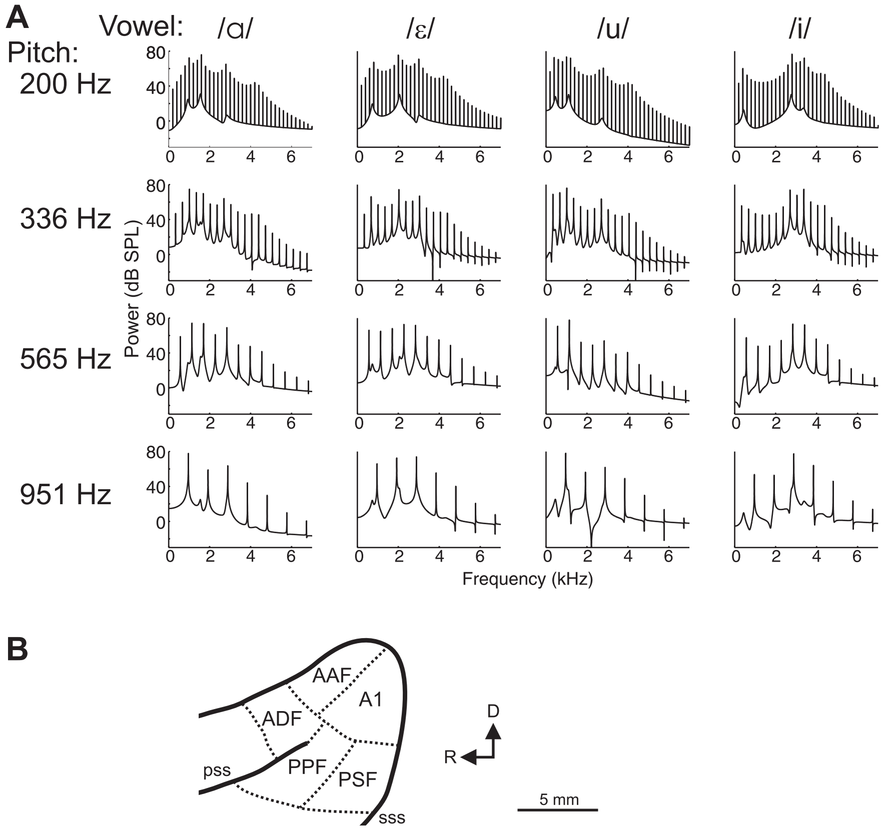

Figure 1. Artificial vowel sounds and locations of five auditory cortical fields from which responses were recorded. ***A***, Frequency spectra of the artificial vowels used in the experiment, fully varied across four spectral timbres (columns) and four pitches (rows). Each of these stimuli was presented at four different virtual locations along the azimuth (−45, −15, 15, and 45° from midline). ***B***, Locations of five auditory cortical fields on the ectosylvian gyrus of the ferret: A1, AAF, PPF, PSF, and ADF. The fields are separated with dotted lines, and the pseudosylvian sulcus (pss) and suprasylvian sulcus (sss) are drawn as solid lines.

coded as having a first-spike latency of 321 ms. Finally, we used Victor's "binless" algorithm to arrive at an optimized and low-dimensional representation of the temporal spiking patterns of a unit (Victor, 2002). Victor's binless method has been previously applied to studies of auditory cortex (Nelken et al., 2005), and details of this algorithm are available in the original source (Victor, 2002).

Bias estimation. In Equation 1, the true probabilities of neural responses, $p(r)$ and $p(x,r)$, are estimated from a limited sample of experimental observations and are inherently biased to overestimate the relations between a stimulus and response (Treves and Panzeri, 1995). Several methods for estimating the bias of MI estimations have been shown to be effective for neural data (Panzeri et al., 2007). We used a method based on bootstrapping a randomly shuffled version of the dataset. In this approach, the set of all spike rate responses were randomly reassigned with replacement to the set of stimulus conditions. The MI of this shuffled dataset was then calculated using Equation 1. This procedure was repeated 500 times, and the median of the resulting MI values was taken as the bias estimate (bias). All MI values reported were bias-corrected by subtracting the value of bias from the MI calculated for the original, unshuffled dataset.

Out of interest, the bias for MI values was also estimated using a less conservative, but more computationally efficient approach (Panzeri et al., 2007) as follows:

$$\text{bias} = \frac{\sum_{x \in X} |R_x - 1| - ||R| - 1|}{2N \ln(2)}, \qquad (2)$$

where $|R_x|$ is the cardinality of responses to stimulus x, $|R|$ is the total number of unique responses observed across all stimuli, and N is the total number of stimulus presentations. The bias values obtained using the random shuffling procedure and from Equation 2 were very similar, and the type of bias estimation used made no difference to the overall pattern of results across fields and stimulus features.

Behavioral testing. Four water-restricted ferrets were trained to detect a change in the pitch or timbre of a repeating artificial vowel on a go/no-go task. The animal initiated each trial by inserting its nose in a poke hole situated at the center of the sound-isolated testing chamber, and a sequence of artificial vowels was then delivered from a speaker positioned above the nose poke hole. Vowels were two-formant versions of those described above, and the vowel /a/ with an F0 of 200 Hz served as the reference sound (80 dB SPL, 350 ms duration, 200 ms interstimulus intervals). The reference sound could change in identity or pitch at the third to seventh vowel in the sequence, and if ferrets withdrew from the nose poke hole during presentation of such a deviant, they were rewarded with water. Failures to withdraw to a deviant (within a 550 ms time window following deviant onset) resulted in a 12 s time out. Spectral timbre and pitch changes were tested in separate sessions, and ferrets completed an average of 93 trials per testing session. Ferrets 1, 2, 3, and 4 completed 12, 22, 25, and 24 sessions on the pitch change detection task and 18, 22, 31, and 41 sessions on the timbre task, respectively. Reaction times were derived from trials in which the animal correctly identified a sound change.

Results

We initially made extracellular recordings from 1083 acoustically responsive units in the left auditory cortex of five ferrets. We characterized the responses of the units to pure tones as well as to artificial vowel stimuli (Fig. 1*A*), which varied along three different parameters: their formant composition (they could be /a/, /ɛ/, /i/, or /u/-like), their fundamental frequency (F0, which is the key determinant of the vowel pitch, could take values of 200, 336, 565, or 951 Hz), and their azimuthal position (which could be −45, −15, 15, or 45°, in which negative values denote directions to the animal's right, contralateral to the recording sites). This therefore produced 64 stimulus combinations. To test whether the units in our dataset were driven by these artificial vowel sounds, we performed paired t tests on the spike rates during vowel presentation (150 ms duration, beginning at sound onset) and during the proceeding silent period (150 ms duration, beginning ≥450 ms after offset of the previous sound). We found that 320 single units and 323 multiunit clusters in our sample were significantly driven by artificial vowels ($p < 0.05$). These units were found in five different areas of auditory cortex (Fig. 1*B*): A1 ($n = 212$ neural units), AAF ($n = 111$), PPF ($n = 155$), PSF ($n = 110$), and ADF ($n = 55$), as judged by their frequency selectivity and temporal response properties. All further analyses were performed on these 643 driven units.

Mutual information throughout the response duration

Using the methodology described below, we calculated the MI between neural spike rates and each of the three perceptual features: timbre, pitch, and azimuth (Eq. 1; see Materials and Methods). In each case, responses were collapsed across the other two

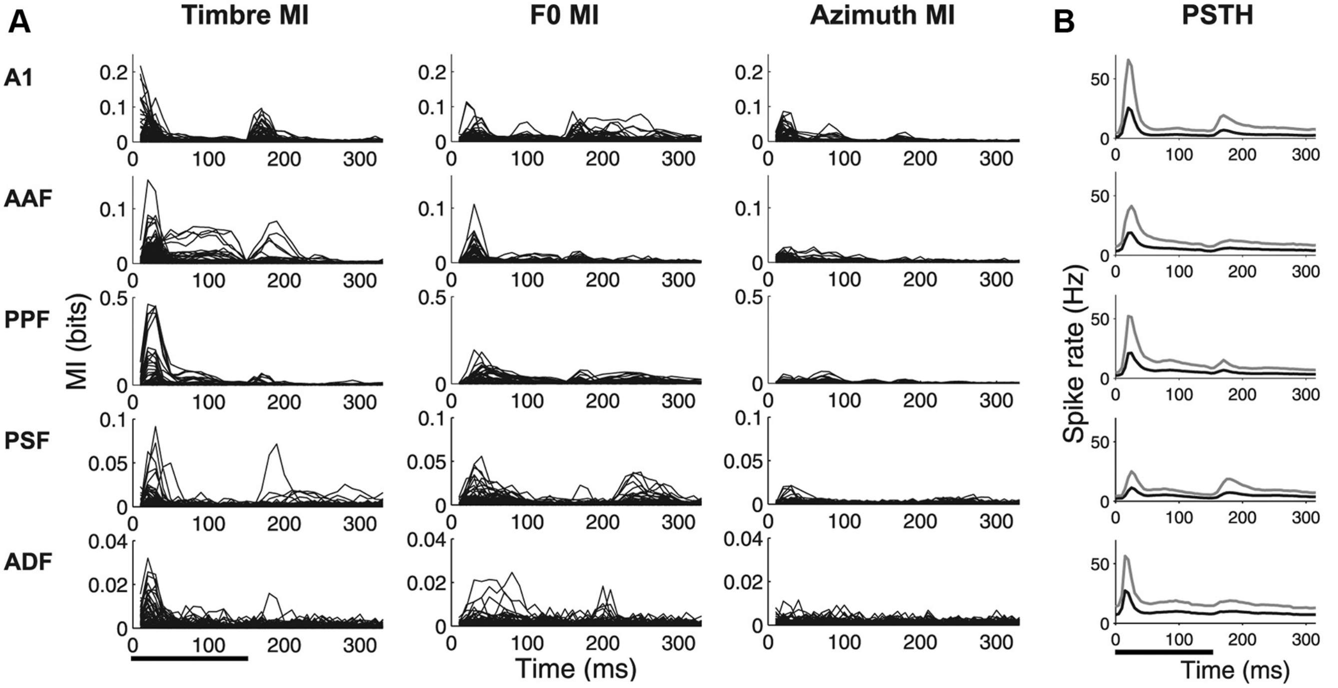

Figure 2. Information for three sound features and average spike rates across the neural response. ***A***, Each plot shows the MI between the relevant sound feature and the spike rate of a different neuron. Information for vowel timbre, periodicity (F0), and azimuth is shown in three columns (left to right, respectively), and neurons from each of the five cortical fields are plotted separately (rows). Each trace indicates, for a given unit, the bias-corrected MI calculated from the spike count within 20-ms-wide time windows positioned in poststimulus time as shown on the *x*-axis. Stimulus presentation time is indicated by the black horizontal bar in the bottom panel. ***B***, Corresponding poststimulus time histograms of responses to artificial vowels. The panels from top to bottom show the average spike rates (mean + SD; black and gray lines, respectively) of units recorded in the five fields (rows).

features. In this way, when we calculate the MI for timbre, we accept the potential confound that arises from allowing pitch and azimuth to vary, so the result reflects timbre information that is robust across pitch and azimuth changes. When choosing a poststimulus time period over which to analyze spike trains, the experimenter usually makes assumptions about the most appropriate temporal response window for the neurons under study. Here, we explored a wide range of possible response windows, repeating MI calculations for spike rates calculated in 20 ms time windows that slide, in 10 ms steps, across the initial 320 ms after stimulus onset. The results are shown in Figure 2*A*. Each line in the MI plots shows the bias-corrected MI for one unit, as a function of poststimulus time. Data are grouped according to which cortical field they were recorded in (rows). The MI estimated for timbre, pitch, and azimuth is displayed in the three columns of Figure 2*A* (left to right, respectively).

Figure 2*A* reveals trends in the distribution of MI across the duration of the response. For instance, the most informative part of the neural response for all three stimulus parameters was often found in the onset component (i.e., during the first 40 ms following stimulus onset), although the sustained and offset responses were also informative in some cortical fields. These plots also show that, while neurons that were most informative about timbre and F0 were found in PPF, the most informative units about azimuth were located in A1. The distributions of stimulus-related information as a function of poststimulus time bear some resemblance to the population-averaged firing rate PSTHs shown in Figure 2*B*. That is, periods of elevated firing rates in a given field tended to be periods of high stimulus-related information. However, there was also considerable unit-to-unit variability in the amount of stimulus-related information signaled and in how that information was distributed over the duration of the response. Informative units were found across a range of BFs in each field (Fig. 3).

To compare the MI values across cortical fields and during different response periods, we calculated the MI for timbre, F0, and azimuth carried by spike rates within four separate 40-ms-wide epochs of the neural response: the onset bin (10–50 ms), a sustained response bin (60–100 ms), the offset bin (160–200 ms), and a late response bin well after stimulus offset (300–340 ms). For each epoch, the median MI for one of the stimulus attributes was calculated across all units in each field. These values are mapped onto their corresponding cortical fields in Figure 4 (gray scale), allowing us to visualize the temporal evolution of information about that attribute across the auditory cortex. This is shown for timbre in the top row of Figure 4, where nearly all the stimulus-related information was found in the onset response, particularly in AAF. In contrast, information about pitch (Fig. 4, middle row) was more distributed across A1, AAF, and the posterior fields during the onset response, and remained high in the posterior fields throughout the sustained response. Pitch information in A1 rose again during the response to vowel offsets and was relatively high in PPF even into the late response epoch. Finally, azimuth information was more confined to the onset responses in the auditory core (Fig. 4, bottom row) and was present to a lesser extent in early PPF responses.

For each of the three stimulus parameters, a Kruskal–Wallis test ($\alpha = 0.05$), with pairwise, *post hoc* multiple comparisons using Tukey's honestly significant difference criterion (Tukey's HSD test) was conducted to examine whether the MI within each response epoch differed across the cortex (Fig. 5). This analysis showed that, for timbre and azimuth, the onset response was generally more informative in the auditory core (A1 and AAF) than in the belt regions studied (PPF, PSF, ADF). The sustained

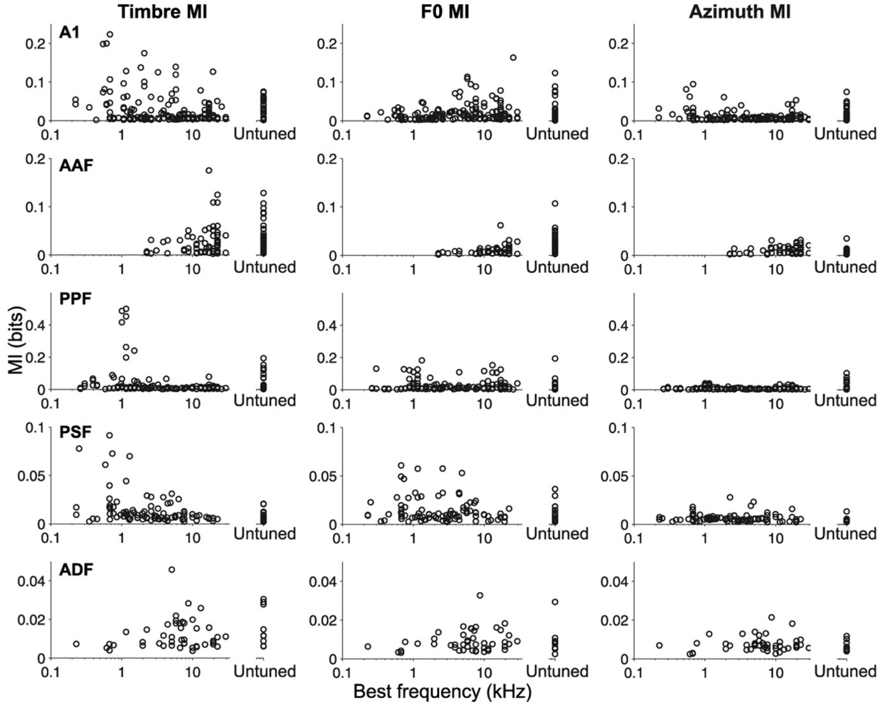

Figure 3. MI as a function of BF. In each scatter plot, the maximum MI value for each unit, across all tested response windows, is plotted (*y*-axis) against the BF for the neuron (*x*-axis). Peak MI values for untuned neurons are plotted to the right of each scatter plot. The 15 plots show data from each of the five cortical fields (rows) and for the three sound features (columns).

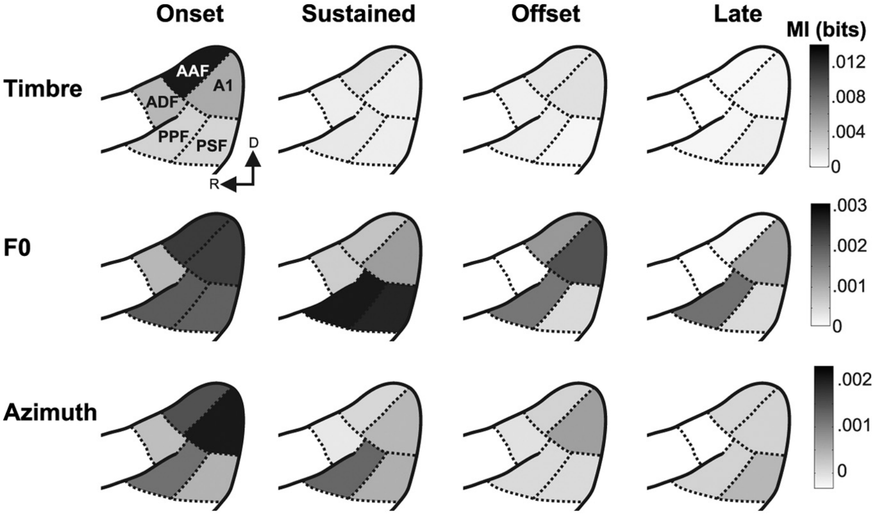

Figure 4. Population median information about timbre, pitch, and location across cortical fields and response periods. In each map of auditory cortex, the gray scale of a cortical field indicates the median MI between a particular stimulus attribute (rows) and spike rates during one of four different 40 ms response window (columns). MI medians were calculated across all units recorded in the field. The gray scales were normalized separately for each stimulus attribute (far right), with darker shades indicating more information.

response was most informative about azimuth in PPF, timbre in AAF and PPF, and about F0 in PPF and PSF. The offset responses of A1 neurons were more informative about timbre and, particularly, F0 than those of neurons in other fields, although offsets in AAF and PPF carried more information about F0 than those in PSF. Finally, the late spike rates of neurons measured 150 ms after vowel offset were more informative for F0 in PPF than in most other fields. This late response was also significantly more informative for timbre in PSF than in the rest of the cortex, although the late MI values for both timbre and azimuth were low overall during this period (note the *y*-axes).

These results show that, for many neurons, the spike rate within one of the four 40 ms response windows described provides information about the pitch, timbre, and azimuth of vowels, but this may not be the most effective way of decoding the spiking responses. We compared the information provided by the best of these four 40 ms response windows for each unit with four other spike codes: a binary spike/no-spike code (Bin), the spike rate calculated over the entire 320 ms duration (R320), the relative first-spike latency (Lat), and Victor's binless algorithm for classifying the temporal spiking patterns (VB). Details of how each code was calculated are provided in Materials and Methods. Codes were compared using a Kruskal–Wallis test with pairwise comparisons using Tukey's HSD tests (Fig. 6). The results showed that for all three stimulus features the binary and full spike count codes provided less information than the other three codes (Tukey's HSD, $p < 0.05$). For pitch and azimuth, the first-spike latency also provided less information than the 40 ms spike rate and Victor's binless classification of the temporal spiking patterns ($p < 0.05$). Therefore, the spike rate within a single, well chosen, and relatively short response window can provide a substantial portion of the information about vowel timbre, pitch, and azimuth that is available in the spike train.

Figure 2*A* shows that, across all fields and response bins, neural spike rates tended to be more informative about the timbre and F0 of vowels than their azimuthal position. A Kruskal–Wallis analysis with pairwise Tukey's HSD tests confirmed that, in many cases, the median MI for F0 and timbre was significantly greater than that for azimuth, particularly for onset and offset responses (data not shown). In contrast, in no cortical field or response window did the median azimuth information exceed that of timbre or periodicity. While the three stimulus parameter ranges were not carefully matched in discriminability, the stimuli presented in this experiment were nevertheless very widely spaced relative to ferret behavioral discrimination thresholds for F0 (Walker et al., 2009) and timbre (our unpublished observations) of artificial vowels, and for the azimuth of noise bursts (Parsons et al., 1999; Nodal et al., 2010). Ferrets can lateralize these artificial vowels (Walker et al., 2009), but their acuity for localizing these sounds in azimuth has not been tested. Given the prevalence of monaural spectral cues and interaural level cues in the high-frequency (>5 kHz) range (Schnupp et al., 2003), one might expect localization accuracy to be better for noise than our low-pass-filtered vowel sounds. We tested this prediction in an additional ferret by comparing the neural responses to our set of vowels with the responses to broadband noises (200 ms duration, 5 ms onset/offset ramps) that varied across four spatial locations (−60, −30, 0, and 30°) using the same VAS filters. In the onset responses of 13 AAF neurons and the sustained responses of 20 neurons in PPF, the MI for azimuth was up to 10 times larger for noise bursts than for vowels (Fig. 7; paired *t* test, $t = 4.82$, $p < 0.05$).

The MI of neural spike rates for vowel features were in the range of ∼0.01–0.6 bits, or ∼1–30% of the total entropy for a given feature, which falls within the range of values reported in previous studies of sensory coding by cortical cells (Panzeri et al., 2001; Mrsic-Flogel et al., 2003; Nelken et al., 2005; Gourévitch and Eggermont, 2010). Although the information provided by any one neuron is insufficient to support highly accurate sensory discrimination, a small number of neurons providing 0.6 bits of independent timbre information each would be sufficient to discriminate the timbre of our sounds to a high level of accuracy, despite the azimuth and pitch confounds. While the onset spike responses of 74% of simultaneously recorded pairs of units in our dataset were significantly correlated (Spearman's correlation; $p < 0.05$), these correlations were generally quite weak (mean $r = 0.12$; SD, 0.12), as previously reported for cortical cells (Brosch and Schreiner, 1999; de la Rocha et al., 2007). Thus, the units we have recorded are largely, although not entirely, independent in their responses, and it is reasonable to expect that a small population of these units could support ferrets' behavioral discrimination of the pitch, timbre, and azimuth of vowels (see behavioral results below).

Tukey's HSD tests further showed that onset responses in A1, AAF, and ADF were, on average, more informative about the timbre of vowels than vowel F0. In contrast, there was more information about sound F0 than about timbre or azimuth in the sustained responses of the posterior fields (PPF and PSF), and in the late responses of A1 and PPF. Thus, while spectral timbre may modulate the onset of cortical responses more strongly, modulation of their spike rates by stimulus pitch seems to persist for longer during the response to a vowel.

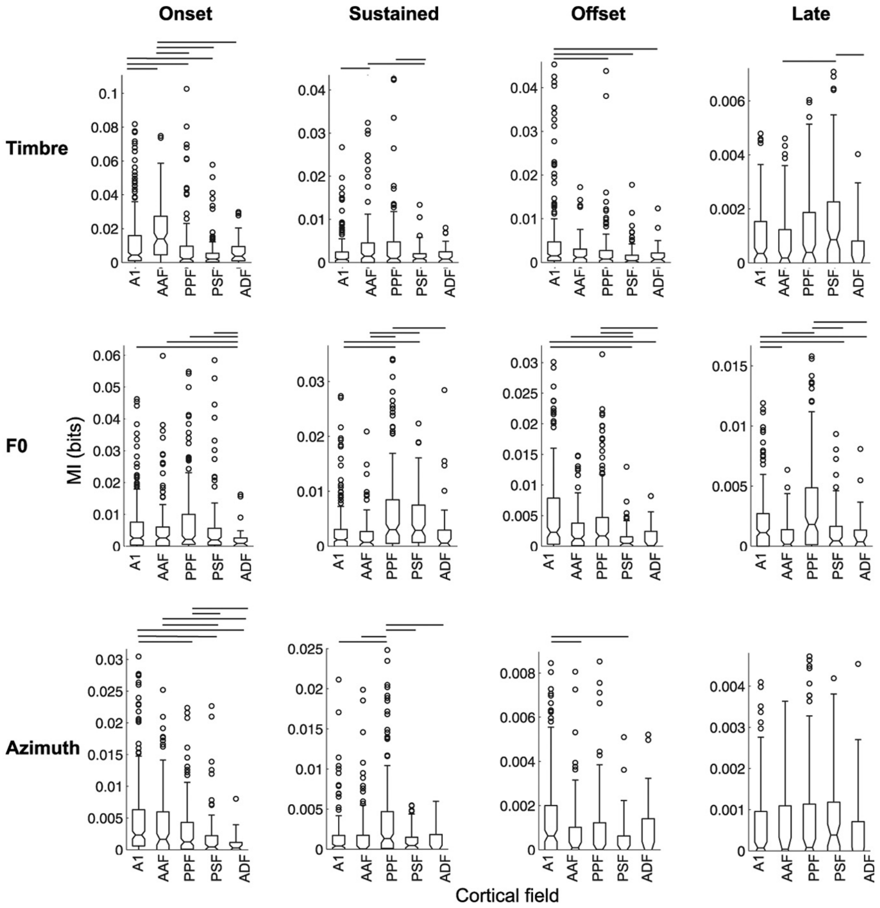

Figure 5. Differences in spike-rate information for timbre, pitch, and location across cortical fields. Box-and-whisker plots show the distribution within the auditory cortex of information about vowel timbre (top row), periodicity (middle row), and azimuth (bottom row) carried by the spike rates of neurons during four different 40 ms response windows (columns). Within each plot, results are grouped according to cortical field. MI values differed significantly across the auditory cortex for all comparisons (Kruskal–Wallis tests, $p < 0.05$), except those of late responses to azimuth (bottom right). The lines at the top of each plot indicate significant pairwise differences in MI between individual fields (Tukey's HSD tests, $p < 0.05$). The ordinate axis is clipped at the 98th percentile for clarity.

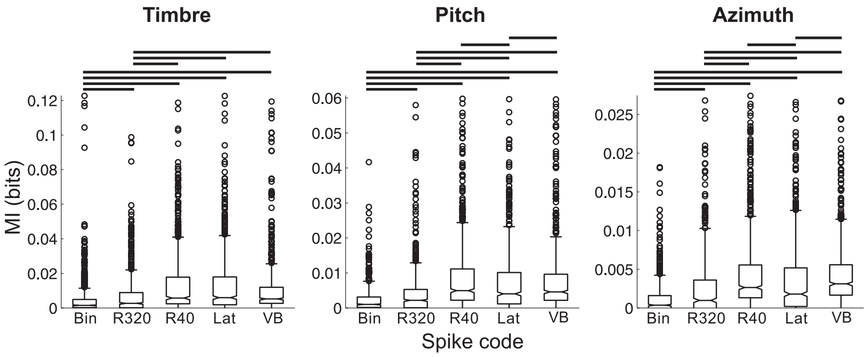

Figure 6. Mutual information across five neural codes. Each panel shows the MI for a given stimulus feature as calculated using each of five spike codes: a binary spike/no-spike (Bin), the overall spike rate (320 ms response period) (R320), the spike rate within the most informative 40 ms response bin (R40), the first spike latency (Lat), and Victor's binless classification of the temporal discharge patterns (VB). Lines above the box-and-whisker plots indicate significant differences (Kruskal–Wallis test, with pairwise Tukey's HSD tests, $p < 0.05$). The ordinate axis is clipped at the 98th percentile for clarity.

These data therefore indicate that sensitivity to pitch, timbre, and location is distributed differentially in the population responses across both time and cortical fields. But does this hold true for individual units? That is, does the same neuron provide information about different stimulus attributes at different times, or do different neurons provide stimulus-related information in each of these time windows? Figure 8 examines this question by comparing MI values obtained from each unit during the onset

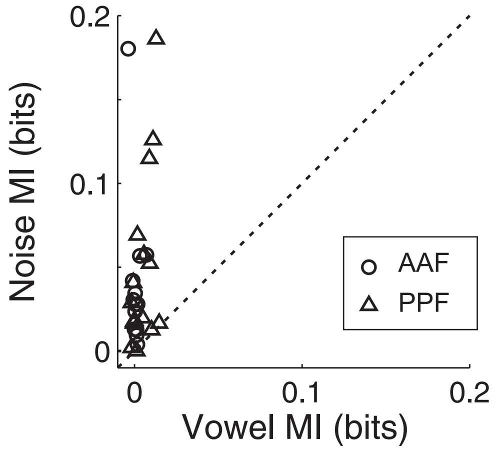

Figure 7. Spike-rate information for the azimuth of broadband noises and artificial vowels. The scatter plot shows the MI of auditory cortical responses for the azimuth of artificial vowels (*x*-axis) and broadband noises (*y*-axis). Units were recorded in AAF (circles; $n = 13$) and PPF (triangles; $n = 20$). The dotted line indicates equality.

response with the MI in the sustained (Fig. 8*A*) or offset responses (Fig. 8*B*). Again, data are plotted separately for each cortical field. This figure shows evidence for large-scale multiplexing of information about the three stimulus features across different time periods within the population of neurons in a cortical field. For example, Figure 8*A* shows that all the units recorded in PPF transmitted more information about timbre (first column) in their onset than in their sustained responses, whereas the opposite was true for azimuth (third column). Because information about a particular stimulus feature could be restricted to particular response periods, this opens up the possibility of multiplexing within the responses of these neurons, with different periods being used to encode different features.

But not all fields were so homogeneous. In some fields, two types of unit were observed: one that primarily encodes the stimulus parameter in the onset response, and another in which MI is much higher during the offset response. An example of this is the encoding of timbre in A1 (Fig. 8*B*, top row, left column). This plot reveals that the most informative neurons fall into two subpopulations: one encodes timbre during onset spike rates, but not during offsets, whereas the second shows the opposite pattern. By contrast, there are very few data points in the top right corner of the scatter plots (i.e., neurons are rarely equally highly informative during onset and offset).

When we examined the time profile of pitch and timbre MI within individual units, we found several examples of neurons that represent these parameters of vowels during separate response windows. This provides more direct support for multiplexing within individual neurons. Examples of multiplexing within single neurons are shown in the seven panels of Figure 9. These examples were chosen to illustrate the variety of ways in which neurons can encode robust information about the timbre and pitch of a vowel by modulating their spike rates in different time windows. The examples shown come from A1, PPF, and PSF, and all were classified as single units.

To estimate what proportion of units exhibit this coding property, we developed an algorithm to identify those that multiplex F0 and timbre in this manner. In this algorithm, we examined the MI for each parameter in individual 20 ms windows across the initial 320 ms of the response (as in Fig. 2*A*). First, we required that the unit in question carried a minimum amount of information about both pitch and timbre. Specifically, the unit had to carry at least 0.01 bits of information about sound timbre in at least one time window examined, and at least 0.01 bits of pitch information in any other window (a criterion of 0.01 bits was based on bootstrapping analyses, described below). Next, we required that, in at least one 20 ms response window, the MI for timbre was >65% of the maximum timbre MI for that neuron, while the pitch MI was <35% of the maximum pitch MI (and thus essentially "invariant" to pitch changes). Finally, we required that the reverse trend was found in another temporal window: pitch MI was >65% of the maximum of the neuron, while timbre MI was <35% of maximum. A unit meeting all of these criteria was classified as "multiplexing." In this manner, a total of 71 units (11% of all units driven by artificial vowels) was classified as multiplexers: 31 in A1 (25 single units and 6 multiunit clusters), 7 in AAF (1 single), 22 in PPF (17 single), 10 in PSF (8 single), and 1 in ADF (0 single).

The highest MI value that an ideal neural response could provide about each of our stimulus features (the "stimulus entropy") is 2 bits, since there were four, equally probable stimulus values for each stimulus parameter (see Materials and Methods). In our dataset, the highest MI value obtained was near 0.5 bits (from a PPF neuron), but for some cortical fields the maximum MI values we estimated were much smaller than this. For example, the most informative unit in PSF provided ~0.1 bits of information about vowel timbre. To determine whether the estimated MI values for any one neuron were greater than might be expected by chance, we used a bootstrapping approach similar to that used in our bias estimation method (see Materials and Methods) to derive a test of the significance of the MI values. If the MI value calculated for a given response window was larger than the 95th percentile of the MI values calculated for scrambled versions of these data (bootstrapped 500 times), the former was considered statistically significant ($p < 0.05$) and we concluded that this unit may constructively contribute to categorizing sounds along the feature dimension in question. This procedure was repeated for 20-ms-wide windows that were positioned from 0 to 700 ms poststimulus offset, in 10 ms steps.

The proportions of significantly informative neurons in each 20 ms poststimulus time window for each stimulus attribute within each cortical field are shown in Figure 10*A*. Both the overall proportion of informative units and the time course of the informative spiking windows varied substantially between fields and stimulus attributes. Common to all fields and stimulus features, however, is that the highest proportions of neurons were informative during the onset response (i.e., during the initial 50 ms). All five fields also share the property that the percentage of timbre-informative units peaked earlier than the percentage of pitch-informative units. In the auditory core (A1 and AAF), the proportion of azimuth-informative units also peaked before pitch-related information. Finally, there was a general tendency for pitch information to persist longer than timbre or azimuth information, particularly in A1 and the posterior fields.

Figure 10*A* emphasizes that the majority of driven units in our sample carried robust information about pitch and timbre at some point in the neural response. Across the five fields, 80% of units carried robust information about vowel timbre in the presence of pitch and azimuth changes (i.e., these units contained significant information in >5% of the time bins examined). By the same metric, 83% of units were robustly informative about pitch, while 61% of units provided robust information about azimuthal location.

Figure 10*A* also reveals differences between cortical fields in the manner in which an individual sound attribute is encoded.

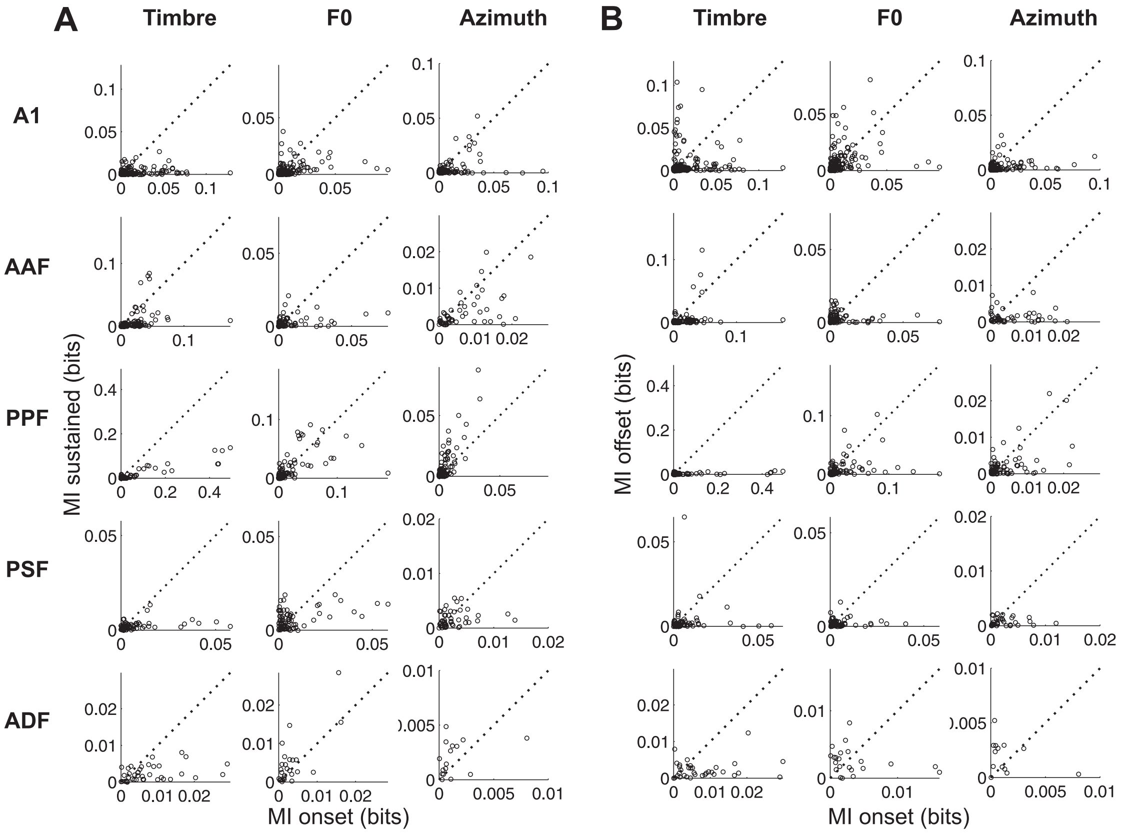

Figure 8. Spike-rate information for timbre, pitch, and location during onset, sustained, and offset responses. ***A***, Scatter plots of the MI based on spike rates for a 40 ms response window during the onset (*x*-axis) and sustained (*y*-axis) period of the neural response. Each circle represents a different unit, with the data plotted separately for units located in the five auditory cortical fields (rows), and for information about the timbre, F0, and azimuth of the sound (columns). The dotted lines are drawn at equality. ***B***, Scatter plots, as in ***A***, comparing the MI in the onset (*x*-axis) and offset (*y*-axis) response windows.

For instance, although timbre information was found predominantly in the onset response in all cortical areas, the sustained portion of the neural response was also informative for timbre in A1, AAF, and PPF, but not in PSF or ADF. Similarly, sustained responses often carried F0 information, but overwhelmingly so in A1 and PPF.

Informative spike count windows tended to correspond to portions of the PSTH where average spike rates were well above the spontaneous level (compare Figs. 2*B*, 10*A*). However, the spike rates of a subset of neurons were informative even at times in the response where the overall spike rate of the population of neurons within the cortical field was near spontaneous firing levels. For example, the average spike rate from 200 to 250 ms poststimulus onset was near the spontaneous rate in all fields, yet a substantial number of units in A1 and PPF encoded significant information about the periodicity of the sound during this time window. This occurs because some units here reliably produced sustained responses to a small subset of the stimulus conditions. Conversely, although we observed a large onset response in ADF, these onset responses were often produced in response to most stimulus conditions and so carried little information about either sound pitch or azimuth.

We next examined the buildup of information about each stimulus feature across the five cortical fields. For each unit, MI was calculated for the spike rate within a time window beginning at sound onset, for durations of 5–150 ms, sampled in 5 ms steps. For the majority of neurons in most fields, the MI increased over the first 40 ms, and then plateaued at some maximum value. The exception was A1, where a sharp peak within the initial 40 ms was typical. To quantify the speed of the buildup of stimulus-related information, we determined the latency at which the MI reached 75% of its maximum for each unit and stimulus attribute. The distributions of this information latency for each of the five cortical fields investigated are shown in Figure 10*B*. Note that information about timbre built up ~10 ms faster than information about pitch in A1, and ~30 ms faster in PPF. Furthermore, the MI for azimuth built up earlier than timbre and pitch MI in AAF neurons, but later than timbre MI in A1 and PPF.

These observations lead to the prediction that an ideal observer should be able to judge the spectral formant structure (and hence the identity) of a vowel slightly faster than the pitch of a vowel. We therefore set out to test whether ferrets make vowel timbre judgments faster than pitch judgments. Four ferrets were trained on a go/no-go change detection task in which they were required to withdraw from a nose poke hole for a water reward when either the pitch or spectral timbre of a repeating artificial vowel changed. The stimuli were two-formant versions of those used in our extracellular recordings with a 200 Hz /a/ as the repeating reference. A two-way ANOVA was used to examine the effects of stimulus attribute (pitch or timbre) and difficulty level (three target pitches or three target vowels) on ferrets' percentage correct scores. Mean percentage correct scores across ferrets are shown by the histograms in Figure 11*A*, along with the performance of each animal. There was no difference in ferrets' percentage correct scores on the pitch and timbre task (two-way ANOVA, $p = 0.210$), or in their performance across different levels of difficulty ($p = 0.170$). Reaction times (i.e., withdrawal times from target onset) were calculated over trials in which fer-

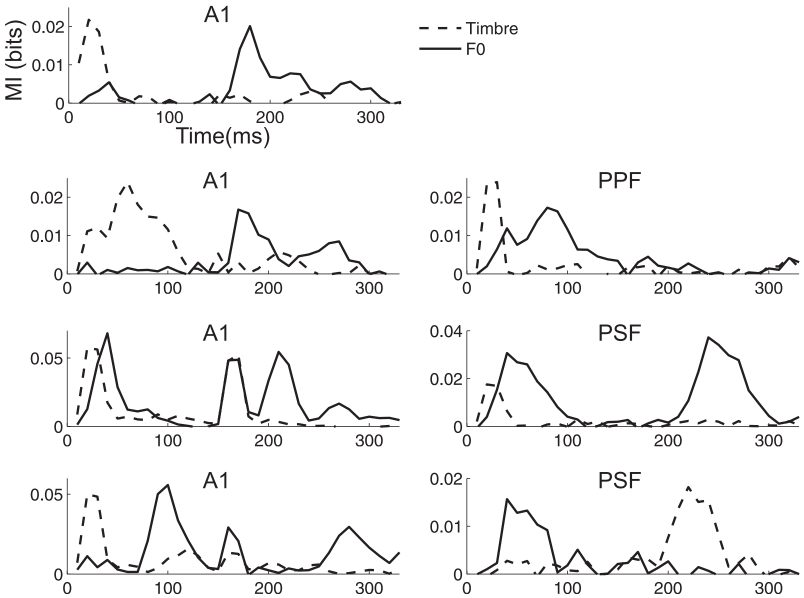

Figure 9. Individual neurons multiplex representations of spectral timbre and F0 in their responses to artificial vowels. Each plot shows MI time traces from a single neuron that was classified as "multiplexing" F0 and timbre information. The robust MI for timbre (dotted line) and periodicity (solid line) were calculated separately in a sliding, 20 ms window. The field from which a given neuron was recorded is indicated at the top of each plot.

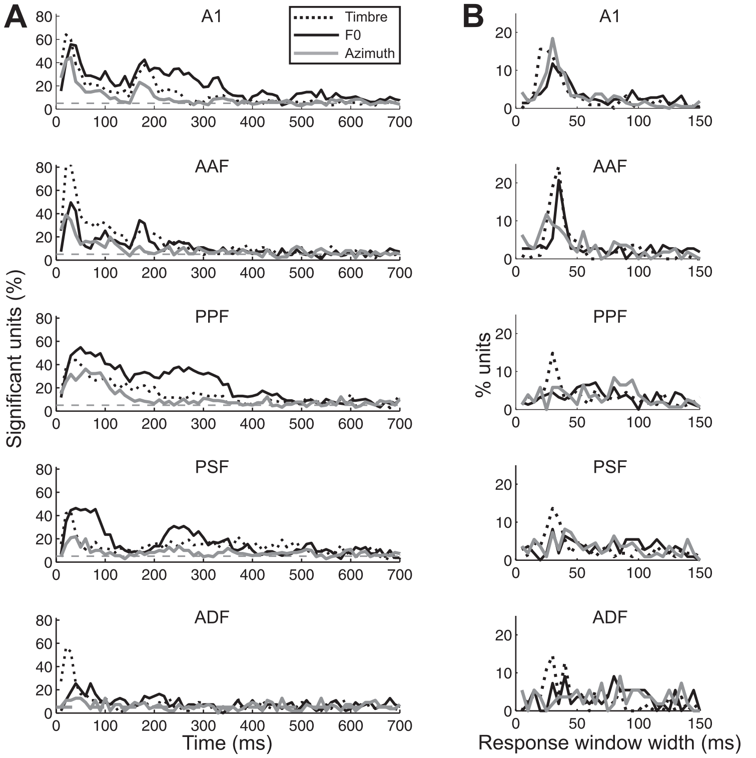

Figure 10. Time profiles of informative neural populations. ***A***, The proportion of units that carried significant information about vowel timbre, F0, and azimuth are plotted across 20 ms time bins (*x*-axis) and the five cortical fields (rows). For each field, the noise in MI values was estimated as the average MI of units during a 20 ms time bin positioned 550 ms after the offset of the sound. The MI for each time bin during the 320 ms response period was thus considered statistically significant if it was larger than the MI calculated for all units during this noise bin. The gray dashed line at 5% indicates the proportion of units that would pass our significance test by chance. ***B***, MI was calculated in time windows positioned at stimulus onset and extending from 5 to 320 ms in duration (tested in 5 ms steps). The response window width at which the MI for a particular unit and stimulus feature reached 75% of its maximum MI value was determined. Histograms of these window widths are shown, across five cortical fields (rows) and for the three stimulus parameters (line styles).

rets made a correct detection. As shown in Figure 11*B*, the ferrets reacted, on average, faster to timbre changes than to pitch changes (two-way ANOVA, $p = 0.007$), with no overall difference across the difficulty levels ($p = 0.120$). Reaction times on the timbre task were ~30 ms faster (mean ± SD, 235.3 ± 26.6 ms) than on the pitch task (264.7 ± 25.8 ms). The reaction time difference observed on these tasks is therefore consistent with our finding that the incidence of significant units (Fig. 10*A*) and MI buildup over the response (Fig. 10*B*) both peak earlier for timbre than for pitch cues in auditory cortex.

Discussion

Our previous work demonstrated that most neurons in ferret auditory cortex are sensitive to multiple perceptual attributes of sound, including the pitch, timbre, and spatial location of the sound (Bizley et al., 2009). This form of "feature combination sensitivity" may help with perceptual binding, whereby several acoustical features must be attributed to a single sound source. However, it raises the problem of perceptual invariance. The melody of a familiar song can be recognized even when it is played on different musical instruments placed at different locations. If cortical neurons are tuned to combinations of features, how can a listener attend to one attribute only, and isolate it from other, irrelevant features? We have shown that, despite the abundance of feature combination sensitivity in auditory cortex, many neurons remain informative about the timbre, pitch, or azimuth of vowels despite large changes in the other two features.

Differences across cortical fields

Although neural responses in AAF carried, on average, more information about the spectral identity of vowels than in other fields, particularly during the onset response, a small group of units in PPF were the most informative about timbre. Pitch information was highest in posterior fields, whereas spatial information was highest in A1, in terms of both the median MI across neurons in the cortical fields and the maximum MI in individual units. The increase in information about spectral timbre and pitch in the posterior bank, relative to A1 and AAF, is analogous to the invariance increase documented within the visual ventral stream (Ison and Quiroga, 2008). Thus, projections from auditory core to the posterior bank of the ectosylvian gyrus in the ferret could mark

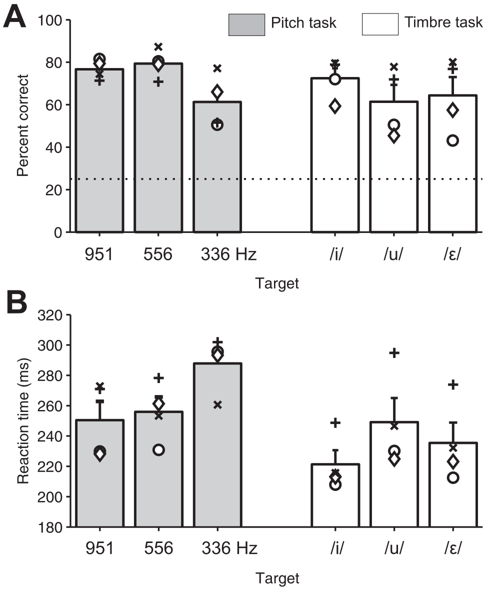

Figure 11. Ability of ferrets to detect changes in the pitch or timbre of artificial vowels. ***A***, Average performance on a go/no-go change detection task across four ferrets (mean bars + SEM), along with the mean score for each ferret (symbols, averaged across sessions). Data are grouped according to the target value, which could change in pitch (gray bars) or timbre (white bars) from a 200 Hz /a/ reference. ***B***, Average withdrawal time following target onset for four ferrets (mean bars + SEM), along with the average withdrawal time for each ferret (symbols). Reaction times were calculated only for correct trials.

the beginning of an auditory stream for object recognition (Rauschecker, 1998; Tian et al., 2001).

We found that cortical fields also differed in the time signatures of their responses. Neurons in the auditory core tended to respond phasically to sound onsets and offsets, while in the belt regions tonic responses were common. Moreover, each of the response windows analyzed could be independently informative about the location, periodicity, and timbre of vowels. Other authors have reported sustained responses in A1, which sometimes lasted well beyond stimulus offset (Moshitch et al., 2006; Campbell et al., 2010). We found that many PPF units carried pitch information up to 200 ms after vowel offset. Therefore, while sustained responses to steady-state sounds may be more common in the auditory cortex of alert animals (Wang et al., 2005; Walker et al., 2008), sustained and informative spiking patterns can also be observed under anesthesia, especially in the higher cortical fields. The medetomidine/ketamine anesthesia used here does not obviously alter the ability of auditory cortical neurons to distinguish natural sounds (Walker et al., 2008), undergo stimulus timing-dependent plasticity (Dahmen et al., 2008), or adapt to sound contrast (Rabinowitz et al., 2011). However, auditory cortical representations can change when animals engage in behavioral tasks (Fritz et al., 2005; Otazu et al., 2009; Lee and Middlebrooks, 2011). Whether task demands affect the manner in which neurons "multiplex" information remains to be characterized in future experiments.

The spiking response of individual units accounted for ~2–30% of the total entropy for any one vowel feature, which is in line with previous findings (Panzeri et al., 2001; Mrsic-Flogel et al., 2003; Nelken et al., 2005; Gourévitch and Eggermont, 2010) and suggests that ferrets' discrimination decisions are likely to be based on small populations of units. Studies that have compared behavioral judgments to neural responses in the auditory (Stecker et al., 2005; Walker et al., 2008; Bizley et al., 2010), visual (Britten et al., 1996), and somatosensory (Hernández et al., 2000) cortices indicate that perceptual decisions likely reflect the activity of populations of neurons, which can carry more stimulus-related information than single neurons alone. Chechik et al. (2006) have shown the information content of auditory cortical neurons to be highly independent, so feature information should improve substantially when the largely uncorrelated responses of cortical neurons are pooled.

Multiplexing neural representations of stimulus features

A novel finding of the present work is the demonstration of feature multiplexing in the auditory cortex. By transmitting multiple signals simultaneously but unambiguously over a shared channel, multiplexing potentially enables individual neurons to encode features along several perceptual dimensions, which may help solve the binding problem while still allowing for perceptual invariance. Other forms of multiplexing in neural responses have been described. For example, in primary visual cortex, stimulus contrast is represented by spike times on a fine timescale (10–30 ms), whereas other visual attributes, such as orientation, spatial frequency, and texture, are represented on coarser (up to ~100 ms) timescales (Victor and Purpura, 1996). There is evidence for this form of multiplexing in our data (Fig. 10*A*), where representations of sound periodicity often persisted over longer durations than timbre responses in A1, PPF, and PSF. In the visual and auditory cortices, multiplexing can occur across even wider timescales, as local field potential oscillations have been shown to carry information that complements the spike-based codes of single neurons (Eggermont and Smith, 1995; Montemurro et al., 2008; Kayser et al., 2009).

Populations of neurons have additionally been shown to represent multiple sound features by using a combination of neural codes. Chase and Young (2005, 2006, 2008) showed that the spike rates of neurons in the inferior colliculus often carried information about more than one type of localization cue, with considerable confounded information between cue combinations (Chase and Young, 2005). By incorporating first-spike latency and spike pattern codes with this spike rate code, these neurons could represent different localization cues more independently (Chase and Young, 2006, 2008).

We provide evidence for an additional type of multiplexing in the responses of auditory cortical neurons. Here, neurons multiplex representations of different sound features by independently modulating their firing rate within discrete time windows throughout their response to a sound. We found that single neurons can thereby provide mutually invariant representations of sound pitch and timbre. Previous studies have shown that specific temporal windows can be informative about the basic properties of sound. Thus, the frequency of pure tones can be encoded in the onset, sustained, and offset responses of cortical neurons (Takahashi et al., 2004; Moshitch et al., 2006; Qin et al., 2007; Fishman and Steinschneider, 2009), and the spectral content of time-varying sounds is represented by spikes evoked throughout the stimulus duration (Qin et al., 2008; Brown and Harrison, 2009). The present finding expands on these results to show that such temporally delimited representations of multiple features can coexist within a single neuron, while remaining essentially invariant to one another.

We found that the neuronal time signature that comprises this form of pitch and timbre multiplexing varies across single neurons and is often temporally complex (Fig. 9). The responses of neurons that multiplex timbre and pitch representations also often contained time windows in which the spike rate was modulated by both these sound attributes. Therefore, upstream neurons could extract either confounded or independent stimulus attributes from this temporally defined spike rate code, depending on their window of integration. The biophysical mechanisms required to read out temporal input sequences have recently been demonstrated within cortical dendrites (Branco et al., 2010). The multiplexed responses described here would predict that upstream cortical neurons should exhibit a range of temporal integration windows. This remains a testable hypothesis for future studies of cortical physiology, particularly with intracellular recording techniques.

Timbre information precedes pitch

At the population level, the time windows in which each feature was represented varied substantially across the five cortical fields examined, but timbre was generally encoded earlier in the onset response than pitch. An analysis of the buildup of information across the response confirmed that, for most neurons, timbre-related information peaked earlier than information about sound pitch. Importantly, this result agrees with our behavioral data, in which ferrets' reaction times in change detection tasks were faster for timbre than for pitch.

Despite this temporal congruence of the behavioral and physiological responses, reaction times on the timbre task were ~30 ms faster than on the pitch task, whereas the latency of timbre information in our recordings was usually ~10–30 ms ahead of pitch information, depending on the cortical field. We also found that pitch information often persisted for longer than timbre information, extending into the sustained and late periods of the response, particularly in the posterior bank. This could delay pitch judgments so that information is consolidated across a longer time window.

Human psychophysical studies have shown that listeners can identify vowels based on only a fraction of a cycle of the F0 of the vowel, whereas reliable pitch judgments require the presentation of approximately four cycles (Gray, 1942; McKeown and Patterson, 1995). This is consistent with the idea that vowel identity can be determined by detecting formant peaks across tonotopically organized frequency channels, and then comparing this activation pattern against memorized spectral templates (Conley and Keilson, 1995). This computation lends itself to parallel spectral processing in the ascending auditory pathway, and one might expect the brain to be able to perform this very quickly. In contrast, the periodicity pitch of a complex sound is likely to be encoded as the autocorrelation of spike times across neurons, which must be calculated over several periods of the waveform (Cariani, 1999). Our data confirm that formant recognition is rapid and precedes pitch recognition in cortical responses as well as behaviorally.

References

Bendor D, Wang X (2010) Neural coding of periodicity in marmoset auditory cortex. J Neurophysiol 103:1809–1822.

Benson DA, Hienz RD, Goldstein MH Jr (1981) Single-unit activity in the auditory cortex of monkeys actively localizing sound sources: spatial tuning and behavioral dependency. Brain Res 219:249–267.

Bizley JK, Nodal FR, Nelken I, King AJ (2005) Functional organization of ferret auditory cortex. Cereb Cortex 15:1637–1653.

Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. J Neurosci 29:2064–2075.

Bizley JK, Walker KM, King AJ, Schnupp JW (2010) Neural ensemble codes for stimulus periodicity in auditory cortex. J Neurosci 30:5078–5091.

Branco T, Clark BA, Häusser M (2010) Dendritic discrimination of temporal input sequences in cortical neurons. Science 329:1671–1675.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 13:87–100.

Brosch M, Schreiner CE (1999) Correlations between neural discharges are related to receptive field properties in cat primary auditory cortex. Eur J Neurosci 11:3517–3530.

Brown TA, Harrison RV (2009) Responses of neurons in chinchilla auditory cortex to frequency-modulated tones. J Neurophysiol 101:2017–2029.

Campbell RA, Schulz AL, King AJ, Schnupp JW (2010) Brief sounds evoke prolonged responses in anesthetized ferret auditory cortex. J Neurophysiol 103:2783–2793.

Cariani P (1999) Temporal coding of periodicity pitch in the auditory system: an overview. Neural Plast 6:147–172.

Chase SM, Young ED (2005) Limited segregation of different types of sound localization information among classes of units in the inferior colliculus. J Neurosci 25:7575–7585.

Chase SM, Young ED (2006) Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. J Neurosci 26:3889–3898.

Chase SM, Young ED (2008) Cues for sound localization are encoded in multiple aspects of spike trains in the inferior colliculus. J Neurophysiol 99:1672–1682.

Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I (2006) Reduction of information redundancy in the ascending auditory pathway. Neuron 51:359–368.

Conley RA, Keilson SE (1995) Rate representation and discriminability of second formant frequencies for /epsilon/-like steady-state vowels in cat auditory nerve. J Acoust Soc Am 98:3223–3234.

Cover TM, Thomas JA (1991) Elements of information theory. New York: Wiley.

Dahmen JC, Hartley DE, King AJ (2008) Stimulus-timing-dependent plasticity of cortical frequency representation. J Neurosci 28:13629–13639.

de la Rocha J, Doiron B, Shea-Brown E, Josić K, Reyes A (2007) Correlation between neural spike trains increases with firing rate. Nature 448:802–806.

Eggermont JJ, Smith GM (1995) Synchrony between single-unit activity and local field potentials in relation to periodicity coding in primary auditory cortex. J Neurophysiol 73:227–245.

Fishman YI, Steinschneider M (2009) Temporally dynamic frequency tuning of population responses in monkey primary auditory cortex. Hear Res 254:64–76.

Fritz JB, Elhilali M, Shamma SA (2005) Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. J Neurosci 25:7623–7635.

Gourévitch B, Eggermont JJ (2010) Maximum decoding abilities of temporal patterns and synchronized firings: application to auditory neurons responding to click trains and amplitude modulated white noise. J Comput Neurosci 29:253–277.

Gray GW (1942) Phonemic microtomy: the mimimum duration of perceptible speech sounds. Speech Monogr 9:75–90.

Hernández A, Zainos A, Romo R (2000) Neuronal correlates of sensory discrimination in the somatosensory cortex. Proc Natl Acad Sci U S A 97:6191–6196.

Ison MJ, Quiroga RQ (2008) Selectivity and invariance for visual object perception. Front Biosci 13:4889–4903.

Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron 61:597–608.

Langner G, Bonke D, Scheich H (1981) Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. Exp Brain Res 43:11–24.

Lee CC, Middlebrooks JC (2011) Auditory cortex spatial sensitivity sharpens during task performance. Nat Neurosci 14:108–114.

Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S (2009) A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. BMC Neurosci 10:81.

McKeown JD, Patterson RD (1995) The time course of auditory segregation: concurrent vowels that vary in duration. J Acoust Soc Am 98:1866–1877.

Middlebrooks JC, Clock AE, Xu L, Green DM (1994) A panoramic code for sound location by cortical neurons. Science 264:842–844.

Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S (2008)

Phase-of-firing coding of natural visual stimuli in primary visual cortex. Curr Biol 18:375–380.

Moshitch D, Las L, Ulanovsky N, Bar-Yosef O, Nelken I (2006) Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. J Neurophysiol 95:3756–3769.

Mrsic-Flogel TD, Schnupp JW, King AJ (2003) Acoustic factors govern developmental sharpening of spatial tuning in the auditory cortex. Nat Neurosci 6:981–988.

Mrsic-Flogel TD, King AJ, Schnupp JW (2005) Encoding of virtual acoustic space stimuli by neurons in ferret primary auditory cortex. J Neurophysiol 93:3489–3503.

Nelken I, Chechik G, Mrsic-Flogel TD, King AJ, Schnupp JW (2005) Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. J Comput Neurosci 19:199–221.

Nodal FR, Keating P, King AJ (2010) Chronic detachable headphones for acoustic stimulation in freely moving animals. J Neurosci Methods 189:44–50.

Ohl FW, Scheich H (1997) Orderly cortical representation of vowels based on formant interaction. Proc Natl Acad Sci U S A 94:9440–9444.

Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an audiotry task suppresses responses in auditory cortex. Nat Neurosci 12:646–654.

Panzeri S, Petersen RS, Schultz SR, Lebedev M, Diamond ME (2001) The role of spike timing in the coding of stimulus location in rat somatosensory cortex. Neuron 29:769–777.

Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. J Neurophysiol 98:1064–1072.

Panzeri S, Brunel N, Logothetis NK, Kayser C (2010) Sensory neural codes using multiplexed temporal scales. Trends Neurosci 33:111–120.

Parsons CH, Lanyon RG, Schnupp JW, King AJ (1999) Effects of altering spectral cues in infancy on horizontal and vertical sound localization by adult ferrets. J Neurophysiol 82:2294–2309.

Phillips DP, Irvine DR (1981) Responses of single neurons in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and responses to intensity. J Neurophysiol 45:48–58.

Qin L, Chimoto S, Sakai M, Wang J, Sato Y (2007) Comparison between offset and onset responses of primary auditory cortex ON-OFF neurons in awake cats. J Neurophysiol 97:3421–3431.

Qin L, Wang JY, Sato Y (2008) Representations of cat meows and human vowels in the primary auditory cortex of awake cats. J Neurophysiol 99:2305–2319.

Rabinowitz NC, Willmore BD, Schnupp JW, King AJ (2011) Contrast gain control in auditory cortex. Neuron 70:1178–1191.

Rauschecker JP (1998) Cortical processing of complex sounds. Curr Opin Neurobiol 8:516–521.

Schnupp JW, Booth J, King AJ (2003) Modeling individual differences in ferret external ear transfer functions. J Acoust Soc Am 113:2021–2030.

Stecker GC, Harrington IA, Middlebrooks JC (2005) Location coding by opponent neural populations in the auditory cortex. PLoS Biol 3:e78.

Takahashi H, Nakao M, Kaga K (2004) Cortical mapping of auditory-evoked offset responses in rats. Neuroreport 15:1565–1569.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. Science 292:290–293.

Treves A, Panzeri S (1995) The upward bias in measures of information derived from limited data samples. Neural Comput 7:399–407.

Victor JD (2002) Binless strategies for estimation of information from neural data. Phys Rev E Stat Nonlin Soft Matter Phys 66:051903.

Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: a metric-space analysis. J Neurophysiol 76:1310–1326.

Walker KM, Ahmed B, Schnupp JW (2008) Linking cortical spike pattern codes to auditory perception. J Cogn Neurosci 20:135–152.

Walker KM, Schnupp JW, Hart-Schnupp SM, King AJ, Bizley JK (2009) Pitch discrimination by ferrets for simple and complex sounds. J Acoust Soc Am 126:1321–1335.

Wang X, Lu T, Snider RK, Liang L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. Nature 435:341–346.