

# A Spike-Timing Code for Discriminating Conspecific Vocalizations in the Thalamocortical System of Anesthetized and Awake Guinea Pigs

Chloé Huetz,<sup>1,2</sup> Bénédicte Philibert,<sup>1,2</sup> and Jean-Marc Edeline<sup>1,2</sup>

<sup>1</sup>Laboratoire de Neurobiologie de l'Apprentissage, de la Mémoire et de la Communication, Unité Mixte de Recherche, Centre National de la Recherche Scientifique 8620, and <sup>2</sup>Laboratoire de Neurobiologie de l'Apprentissage, de la Mémoire et de la Communication, Université Paris-Sud, 91405 Orsay, France

Understanding how communication sounds are processed and encoded in the central auditory system is critical to understanding the neural bases of acoustic communication. Here, we examined neuronal representations of species-specific vocalizations, which are communication sounds that many species rely on for survival and social interaction. In some species, the evoked responses of auditory cortex neurons are stronger in response to natural conspecific vocalizations than to their time-reversed, spectrally identical, counterparts. We applied information theory-based analyses to single-unit spike trains collected in the auditory cortex ( $n = 139$ ) and auditory thalamus ( $n = 135$ ) of anesthetized animals as well as in the auditory cortex ( $n = 119$ ) of awake guinea pigs during presentation of four conspecific vocalizations. Few thalamic and cortical cells ( $< 10\%$ ) displayed a firing rate preference for the natural version of these vocalizations. In contrast, when the information transmitted by the spike trains was quantified with a temporal precision of 10–50 ms, many cells ( $> 75\%$ ) displayed a significant amount of information (i.e.,  $> 2SD$  above chance levels), especially in the awake condition. The computed correlation index between spike trains ( $R_{\text{corr}}$ , defined by Schreiber et al., 2003) indicated similar spike-timing reliability for both the natural and time-reversed versions of each vocalization, but higher reliability for awake animals compared with anesthetized animals. Based on temporal discharge patterns, even cells that were only weakly responsive to vocalizations displayed a significant level of information. These findings emphasize the importance of temporal discharge patterns as a coding mechanism for natural communication sounds, particularly in awake animals.

**Key words:** auditory cortex; thalamus; spike trains; vocalization; information theory; temporal coding

## Introduction

Acoustic communication is crucial for survival and social interactions in many species. Understanding how conspecific vocalizations are encoded in the auditory pathway is a prerequisite for deciphering the neural bases of acoustic communication. The idea that neuronal specialization has emerged for processing conspecific vocalizations is supported by cumulative evidence indicating that during development the cortical map and functional properties of auditory cortex (ACx) neurons undergo continuous changes to adapt to environmental acoustic stimuli (Chang and Merzenich, 2003; Engineer et al., 2004; Nakahara et al., 2004; Chang et al., 2005; de Villers-Sidani et al., 2007). These develop-

mental dynamics potentially explain why neurons exhibit a strong preference for natural acoustic stimuli present in the animal environment. In songbird, in particular brain areas such as the nucleus HVC (but also the nucleus NIf), neurons display highly selective responses, firing more to playback of the bird's own song (BOS) than to the reversed BOS or other conspecific songs (for review, see Margoliash, 1997; Doupe and Kuhl, 1999; Nealen and Schmidt, 2002; Prather and Mooney, 2004). No such selectivity has been reported in mammals, but, using the firing rate of cortical cells as a metric, most ACx neurons in the marmoset auditory cortex prefer the natural version of the Twitter call over its time-reversed version (Wang et al., 1995; Wang, 2000). This is not the case for neurons recorded in the cat ACx (Wang and Kadia, 2001). The preference for natural versions of conspecific vocalizations over their time-reversed versions is not observed in either the squirrel monkey ACx (Glass and Wollberg, 1983a,b; Pelleg-Toiba and Wollberg, 1991) or cat ACx (Gehr et al., 2000). This discrepancy may depend on several factors, such as the species, type of natural stimuli, and potential sample bias of the type of cortical cells recorded. Recent findings suggest an alternative hypothesis. Schnupp et al. (2006) demonstrated that reliable discriminations between natural and time-reversed vocalizations could be achieved based on temporal discharge patterns operating at the 10 ms scale, indicating that natural stimuli

Received July 14, 2008; revised Oct. 26, 2008; accepted Nov. 18, 2008.

This work was partially supported by Grant ANRNeuro2006 to J.-M.E. C.H. was initially supported by a doctoral fellowship from the Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche and subsequently by an Attaché Temporaire d'Enseignement et de Recherche position from Paris-Sud University. B.P. was supported by a postdoctoral fellowship from Centre National de la Recherche Scientifique. We thank Dr. Mounya Elhilali and Dr. Elizabeth Hennevin for insightful and detailed comments on previous versions of this manuscript. We are very grateful to Nathalie Samson and Pascale Leblanc-Veyrac for taking care of the guinea pig colony. We thank SciTechEdit International for help in improving the language.

Correspondence should be addressed to Dr. Jean-Marc Edeline, Laboratoire de Neurobiologie de l'Apprentissage, de la Mémoire et de la Communication, Unité Mixte de Recherche 8620, Université Paris-Sud, Bât 446, 91405 Orsay, France. E-mail: jean-marc.edeline@u-psud.fr.

DOI:10.1523/JNEUROSCI.3269-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/290334-17\$15.00/0

can be encoded by temporal information rather than by the average firing rate. The aim of the present study was twofold. The first goal was to evaluate whether cortical and thalamic neurons respond with higher rates to conspecific guinea-pig vocalizations compared with those same vocalizations played in reverse (these time-reversed vocalizations have the same frequency power spectrum but the spectrotemporal structure of the natural vocalization is disrupted). The second goal was to evaluate the respective contributions of firing rate and spike timing to the information carried by the natural and time-reversed versions of these vocalizations. The firing rate of auditory thalamus and ACx neurons was first quantified using the index previously used to reveal the selectivity of ACx neurons for the natural version of marmoset vocalizations (Wang and Kadia, 2001). Then, the difference in the temporal organization of spike trains elicited on presentation of the natural and time-reversed vocalizations was quantified using the metric-space method developed by Victor and Purpura (Victor and Purpura, 1996, 1997). This method can be used to determine the relative contribution of the firing rate and the spike-timing precision in the information conveyed by neuronal spike trains. To determine whether spike timing is more reliable on presentation of the natural or time-reversed stimuli, a correlation index was computed ( $R_{\text{corr}}$ ) (Schreiber et al., 2003).

The response selectivity described in animals under anesthesia cannot necessarily be reproduced in awake animals. For example, in the avian brain, the selective responses for the BOS observed during periods of synchronized EEG (anesthesia or slow-wave sleep) were not observed in awake birds (Schmidt and Konishi, 1998; Cardin and Schmidt, 2003, 2004). Similarly, the strength of evoked responses and the temporal discharge patterns in the mammalian auditory system differ considerably between anesthetized and awake animals (Tortorolo et al., 2002; Cotillon-Williams and Edeline, 2003, 2004; Massaux et al., 2004; Populin, 2005); for other references see (Hennevin et al., 2007). Thus, in the present study, to ensure that the anesthetic did not mask some aspect of the neural code that is in use in awake animals, we also analyzed neuronal spike trains recorded in the auditory cortex of awake guinea pigs. Parts of these results were presented in an abstract format (Huetz et al., 2007).

## Materials and Methods

### *Animal preparation and recording procedures*

**Anesthetized animals.** Experiments were performed on 26 adult guinea pigs (390–650 g; national authorization No. 91-271 to conduct animal research) anesthetized by an initial injection of diazepam (6 mg/kg, i.p.) followed by urethane (1.2 g/kg, i.p.). Additional doses of urethane (0.5 g/kg, i.p.) were systematically delivered when reflex movements were observed after pinching the hindpaw (usually once or twice during a given recording session). The body temperature was maintained around 37°C by a heating pad throughout the experiment. The trachea was cannulated and a local anesthetic (xylocaine, 2%) was infiltrated into the wound. The stereotaxic frame supporting the animal was placed in a sound-attenuating chamber (IAC, model AC2).

For the thalamic recordings, a circular hole was drilled in the skull above the medial geniculate body (MGB) and electrodes were vertically penetrated. For the cortical recordings, a large opening was made in the temporal bone and very small slits were made in the dura mater under microscopic control. A diagram of the pattern of vasculature was drawn and the location of the primary field (AI) was estimated based on our previous studies (Edeline et al., 1993; Manunta and Edeline, 1999). The cortical surface was rapidly mapped to confirm the location of AI: neuronal clusters were recorded with low impedance (<1 M $\Omega$ ) electrodes until a progression from low to high frequency was observed in the caudorostral direction (Wallace et al., 2000). During each recording session, the first electrode penetration was made with tungsten microelec-

trodes (>8 M $\Omega$ ) and the following penetrations (at nearby locations) were made with glass micropipettes (5–10 M $\Omega$ ). The signal from the electrode was amplified (gain 1000, bandpass 0.3–10 kHz), and multiplexed to an audio monitor and a voltage window discriminator. The waveform of the action potentials and the corresponding TTL pulses generated by the discriminator were continuously displayed, digitized (50 kHz sampling rate, Superscope, GW Instruments), and stored for off-line analyses. The pulses were sent to the acquisition board (PCLab, PCL 720) of a laboratory microcomputer, which recorded them with a 50  $\mu$ s resolution and provided on-line displays of the neuronal responses. Both in thalamus and cortex, successive recording sites were separated by at least 100  $\mu$ m in depth. At the end of the recording session (10–12 h), the animal was killed by a lethal dose of pentobarbital (200 mg/kg).

**Awake animals.** Adult guinea pigs (390–550 g,  $n = 12$ ) underwent surgery under anesthesia (atropine 0.08 mg/kg, diazepam 8 mg/kg, pentobarbital 20 mg/kg; see (Evans, 1979)). Three silverball electrodes were inserted between the bone and dura: one was used as a reference during the recording sessions; the other two, placed over the frontal and parietal cortices, served to monitor the cortical electroencephalogram (EEG). An array of 5–8 tungsten electrodes ( $\sim$ 1.0 M $\Omega$  at 1 kHz, spaced 200–300  $\mu$ m in the rostrocaudal axis) was slowly inserted in the auditory cortex under electrophysiologic control. Starting from 600  $\mu$ m below the pia, responses to pure tone frequencies were tested at regular depths to optimize the strength of evoked responses; the final placement depth of the electrodes ranged from 800 to 1250  $\mu$ m. A dental acrylic cement pedestal, including two cylindrical threaded tubes, was built to allow for atraumatic fixation of the animal's head during the subsequent recording sessions. An antiseptic ointment (Cidermex, neomycine sulfate, Rhone-Poulenc Rorer) was liberally applied to the wound around the pedestal. All surgical procedures were performed in compliance with the guidelines determined by the national (JO 887–848) and European (86/609/EEC) legislations on animal experimentation, which are similar to those described in the *Guidelines for the Use of Animals in Neuroscience Research of the Society of Neuroscience*. In addition, regular inspections of our laboratory by accredited veterinarians designated by Paris-Sud University confirmed that care was taken to maximize the animals' health and comfort throughout the different phases of the experiment.

Three days after surgery, each animal was adapted to restrained conditions in an acoustically isolated chamber (IAC, model AC2) for several days. The animal was placed in a hammock with the head fixed for increasing periods of time (1–3 h/d). The animal was also accustomed to hearing sequences of pure tone bursts as well as different vocalizations used subsequently to test the neuronal responses.

All animals used in the present study (both in the anesthetized and awake conditions) were housed in a colony room and were grouped by four or five animals in large plastic cages (75  $\times$  55  $\times$  25 cm; Tecniplast, Buguggiate, Italy) with large wire mesh doors (55  $\times$  20 cm). All animals frequently emitted vocalizations during social interactions with the other animals of the same cage and loudly vocalized during animal care and feeding.

### *Auditory stimuli and experimental protocol*

All the cells included in the present study exhibited reliable tuning curves when tested with pure tone frequencies. The sound generating system to deliver pure tone frequencies was the same as that previously described (Edeline et al., 2000, 2001; Manunta and Edeline, 2004). Pure tones (100 ms, rise/fall time 5 ms) were generated by a remotely controlled wave analyzer (Hewlett-Packard model HP 8903B) and attenuated by a passive programmable attenuator (Wavetek, P557, maximal attenuation 127 dB), both controlled via an IEEE bus. Stimuli were delivered through a calibrated earphone (Beyer DT48) placed close to the ear canal. The system was calibrated using a sound level calibrator and a condenser microphone/preamplifier (Bruel and Kjaer models 4133 and 2639T) placed at the same distance from the speaker as the animal's ear (< 5 mm). The whole sound delivery system (HP 8903B, attenuators, and speaker) was calibrated from 0.1–35 kHz and could deliver tones of 80 dB SPL up to 20 kHz and of 70 dB SPL up to 35 kHz. Harmonic distortion products were measured to be down  $\sim$ 50 dB from the fundamental.

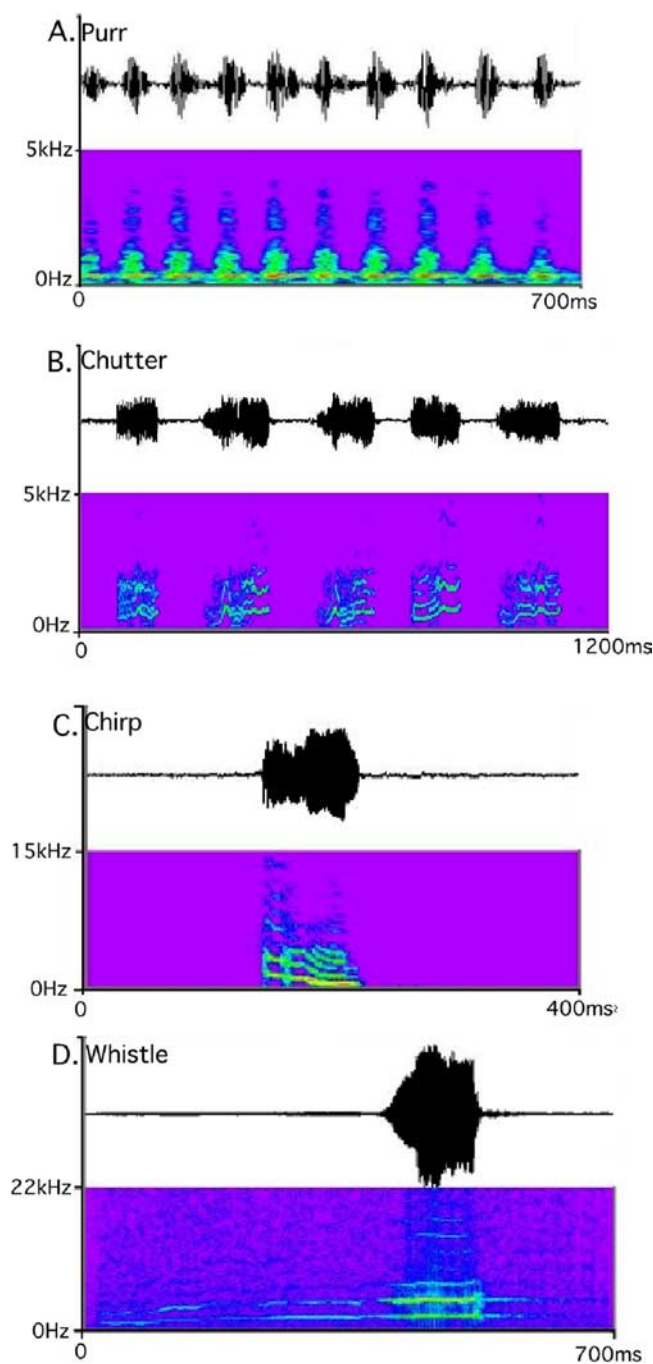
Once the frequency tuning was established, four vocalizations used in

a previous study (Philibert et al., 2005) were presented in their natural and time-reversed versions. These vocalizations were collected from five adult male guinea pigs recorded either in pairs or individually in a sound-attenuated room. Calls were recorded using a Sennheiser MD46 microphone connected to a microcomputer and digitized using SoundEdit software (44 kHz sampling rate). The relationships between these calls and the animal behavioral repertoire have been previously described (Berryman, 1976; Harper, 1976). A “purr” consists of a series of low-frequency impulses (fundamental frequency <500 Hz, duration 700 ms) emitted when social contact is allowed or sought. A “chirp” is a brief call (0.7–15 kHz, <100 ms) that is believed to be a low-intensity distress call or a warning signal. A “chutter” consists of a chain of five components (0.5–3.5 kHz, 150–250 ms separated from each other by 140–175 ms) emitted during discomfort. A “whistle” is a two-part call (250–400 ms, with the first part from 1 to 3 kHz and the second part rising steeply to 8–20 kHz) emitted when animals are isolated or in response to stimuli associated with caretaking. We initially recorded a large set of samples for each vocalization class, and then selected one sample in each vocalization class that we considered to be the most representative of the class. Figure 1 displays the spectrograms and oscillograms of the four selected vocalizations that were used to test all neurons recorded in the present study. The time-reversed versions of the stimuli were generated by reversing the natural calls in the time domain, i.e., playing the call backward. Each call was presented at a peak intensity of 70 dB SPL sound pressure level. The natural and time-reversed versions of the four calls were presented in random order, with each call repeated 20 times with a 2 s period of silence between each vocalization. The whole protocol, i.e., testing the frequency tuning with pure tones and the responses to the four vocalizations (natural and time-reversed) lasted ~50 min. When recording in unanesthetized animals, the EEG was displayed on a polygraph and a computer to make sure that the animal was awake during the entire recording session (the data collection was stopped when large voltage EEG signals characteristic of slow-wave sleep were present). Each unanesthetized animal was recorded during 3–4 recording sessions, separated by 24 or 48 h. In all cases, the recording session was stopped each time the spike waveform became unstable. Systematic off-line examination of the digitized waveforms confirmed that spike trains of unambiguously isolated single units were recorded both in anesthetized and awake animals.

#### Quantification of tuning curves and responses to vocalizations based on firing rate

For each cell, the frequency tuning was quantified from the threshold up to 70 dB SPL by 10 dB steps. At each intensity, the best frequency was determined as the frequency eliciting the largest evoked responses. The breadth of tuning was quantified both by the  $Q_{20\text{ dB}}$  and by the square root transformation  $\sqrt{f_2 - f_1}$ , where  $f_2$  and  $f_1$  indicate the high and low limits of the tuning bandwidth at 20 dB above threshold (this later measure is independent of the unit characteristic frequency (CF) (Whitfield, 1968; Whitfield and Purser, 1972; Calford et al., 1983). The latency of the tone-evoked responses was computed at each intensity used to test the frequency tuning curve. At a given intensity, the responses obtained for all the tested frequencies were considered, and the latency of the first spike after tone onset was computed (1 ms precision). For each cell and at each intensity, the variability of the latency was quantified by the SD of the mean latency value.

Statistical tests (paired *t* test) were used to determine whether a given cell was “responsive” to a particular vocalization. In the case of phasic responses followed by prolonged inhibition, the evoked firing rate did not significantly differ from spontaneous activity. In these cases, a cell was classified as “responsive” to a particular vocalization if evoked responses were obtained for at least 10 of 20 presentations based on the rasters and histograms (5 ms bin), as evaluated by two of the authors. As previously described (Philibert et al., 2005), responsive cells can display different types of responses to vocalizations: (1) phasic (onset or offset) responses correspond to transient evoked discharges at the onset and/or the offset of the vocalizations, (2) phase-locked responses correspond to evoked discharges exhibiting multiple peaks during the purr and/or the chutter, and (3) sustained responses correspond to evoked firing rates above spontaneous activity occurring throughout the vocalization with-



**Figure 1.** Oscillogram (top) and spectrogram (bottom) of the four vocalizations used in the present study.

out noticeable temporal organization. Nonresponsive cells do not exhibit a significantly increased firing rate on stimulus presentation and sometimes show transient inhibition.

According to the type of evoked response, either the first 100 ms (for the phasic responses) or the entire duration of the vocalization (for the phase-locked and sustained responses) was used to quantify the response strength to the natural and time-reversed version of each vocalization. The preference for the natural versus the time-reversed version was quantified by the directional index described by Wang and Kadia (2001):

$$d = (R_{\text{nat}} - R_{\text{rev}}) / (R_{\text{nat}} + R_{\text{rev}}),$$

where  $R_{\text{nat}}$  and  $R_{\text{rev}}$  correspond to the firing rate on presentation of the natural and reversed calls, respectively. A  $d$  value of 1 (–1) indicates that

a neuron responded to only the natural (reversed) call; zero means no firing-rate difference between the natural and reversed calls.

### Spike train analysis with the metric-space method

**General overview.** In a recent study (Huetz et al., 2006), we extensively explained how the method developed by Victor and Purpura (1996, 1997) can be applied to determine whether, in addition to information carried by spike count, the precise timing of the spikes conveys information about acoustic stimuli. Only the major features of this method will be presented here. To consider the whole response pattern, note that only spike trains of equal duration can be processed. Therefore, we only compared the spike trains to a particular vocalization with the spike trains from the time-reversed version of this particular vocalization (and not with spike trains obtained by other vocalizations of different durations).

**Methodological considerations.** The metric-space method involves two stages: (1) computing pairwise distances between spike trains (using the  $D_{\text{count}}$  and  $D_{\text{spike}}$  distance as defined by Victor and Purpura, 1996), and (2) measuring the information ( $H$ ) between stimuli and spike trains as a measure of spike train clustering. The distance between two spike trains taken from two sets of trials (natural and time-reversed vocalization) is defined as the minimal “cost” for transforming one spike train into the other via a sequence of elementary steps. The following elementary steps are allowed: inserting a spike, deleting a spike, and shifting a spike by an amount of time  $dt$ . Each elementary step is associated with a cost. For eliminating or inserting a spike, the cost is unity. To shift a spike, the cost is equal to  $qdt$ , where  $dt$  is the extent of the shift expressed in units of time bins. The parameter  $q$  (in  $\text{s}^{-1}$ ) accounts for and quantifies the temporal precision relevant to spike timing. Pairwise distances between all responses are computed for a given value of  $q$  and clustering is performed.

Clusters can be defined as sets of spike trains that are close to each other. To evaluate the extent of the similarity between spike trains elicited by the same stimulus and those elicited by another stimulus, a confusion matrix  $N(s_p, r_j)$  is constructed. This matrix summarizes, for each stimulus class, how many spike trains can be attributed to this class, based on the average distance (or similarity) of this spike train from other spike trains of the same stimulus class. Here, the confusion matrix is  $2 \times 2$ , where the two columns correspond to the two stimulus classes (natural and time-reversed vocalizations), and the two rows correspond to the two response classes (natural and time-reversed vocalizations). This matrix  $N$  is then used to compute the amount of transmitted information  $H$  (Victor and Purpura, 1996):

$$H = \frac{1}{N_{\text{total}}} \sum_{ij} [N(s_i, r_j) (\log N(s_i, r_j) - \log \sum_a N(s_a, r_j) - \log \sum_b N(s_i, r_b) + \log N_{\text{total}})],$$

in which  $N_{\text{total}}$  is the total number of spike trains (natural and time-reversed vocalizations) and logarithms are in base 2. A perfect clustering of spike trains corresponds to a purely diagonal confusion matrix, and to an information value of  $H = 1$ . In contrast, when clustering is totally random, on and off diagonal values of the matrix are similar, leading to an information value of  $H = 0$ . In such a case, classification is impossible, meaning that either the responses were not correlated with the stimuli, or that the distance (i.e., measure of similarity) used was not appropriate for classifying the spike trains into two categories.

**Classification of the responses as “rate” or “temporal.”** By varying the parameter  $q$ , it is possible to determine which metric  $D_{\text{spike}}[q]$  produces the best clustering of spike trains and thus provides the maximum information value. When  $q = 0$ , distances between spike trains only depend on their spike count (and refers to the  $D_{\text{count}}$  distance in the work of Victor and Purpura, 1996); and as  $q$  increases, distances become sensitive to finer temporal features. Therefore, if information  $H$  is maximized for  $q = 0$ , this indicates that responses to natural versus time-reversed vocalizations can be discriminated exclusively based on spike count. In the following, such cases will be called “rate responses.” Conversely, if the maximum value of  $H$  is attained for  $q > 0$ , this indicates that spike timing is crucial to accurately discriminate responses to natural versus time-reversed vocalizations. In the following, such cases will be called “temporal responses.”

**Computation of the bias.** Amounts of information computed from experimental data are biased estimates of the “true” transmitted information that would ideally be obtained from an infinite number of trials for each stimulus. Therefore, additional computations should be performed to estimate the bias (or chance level) and to assess the significance of transmitted information resulting from the calculation described above. Here, the bias (or chance level)  $H_{\text{bias}}$ , is estimated as described by Victor and Purpura (1997) and as implemented in the Spike Train Analysis Toolkit (<http://neuroanalysis.org/toolkit/>). Information is recalculated after random reassignments of the responses across stimuli: From the set of 20 responses to natural and 20 responses to time-reversed vocalization, 20 new sets of spike trains are formed, in which responses are assigned to a randomly chosen stimulus (natural or time-reversed vocalization). For each set, information values are computed using the same procedure as for normal calculations. The mean and the SD of these 20 values are computed. The mean,  $H_{\text{bias}}$ , represents the level of chance or bias. In reporting the results, only values of  $H$  that were greater than  $H_{\text{bias}} + 2 \text{ SD}$  were considered significant and therefore only those cells that had significantly positive values of  $H$  were included.

**Determination of  $H^*_{\text{max}}$ ,  $H^*_{\text{count}}$ , and  $q^*_{\text{max}}$ .** For each cell, the metric-space analysis and the calculation of the bias were performed for values of  $q$  ranging from  $q = 0$  to  $q = 10^{-5} \text{ s}^{-1}$  in logarithmic steps, thus providing a curve of the transmitted information ( $H$ ) as a function of  $q$  as well as a curve of the bias ( $H_{\text{bias}}$ ) as a function of  $q$ . Three indices were derived from these curves:  $H_{\text{max}}$ ,  $H_{\text{count}}$ , and  $q_{\text{max}}$ .  $H_{\text{max}}$  represents the maximum amount of information that a cell conveys about the stimuli: it quantifies how well stimuli can be discriminated by looking at the spike trains. The index  $q_{\text{max}}$  is the value of  $q$  corresponding to  $H_{\text{max}}$  and therefore the temporal precision with which the cell conveys a maximum amount of information.  $H_{\text{count}}$  represents the amount of information that the spike count of a cell’s responses conveys about the stimuli: it is the value of  $H$  when  $q = 0$ . In Results, we adopted the following notation:  $H^*_{\text{max}}$  and  $H^*_{\text{count}}$  are the values of  $H_{\text{max}}$  and  $H_{\text{count}}$  obtained after subtracting the mean bias value  $H_{\text{bias}}$ , and  $q^*_{\text{max}}$  is the value of  $q$  corresponding to  $H^*_{\text{max}}$ .

### Are temporal discharge patterns observed on presentation of the reverse vocalization “mirror images” of those elicited by the forward vocalization presentations?

The strong trial-to-trial reliability of spike trains observed on some rasters suggests that thalamocortical neurons might be entirely driven by some simple acoustic component in the stimuli. Indeed, if neurons of the auditory thalamocortical system act as simple linear spectral filters with monotonic behavior, responses to time-reversed versions of the vocalizations would correspond to the mirror image of the responses to the natural versions shifted in time by the latency of the neuronal response. In this case, the temporal discharge patterns that emerge during presentation of natural and time-reversed vocalizations should be “mirror images” of each other.

To investigate this hypothesis, we formed a new set of spike trains. Responses to natural vocalizations were kept identical, whereas responses to reversed vocalizations were time-reversed and shifted by an appropriate temporal delay (see below). If responses to reversed vocalizations are simply the mirror image of natural vocalization responses, then the time-reversed versions of spike trains obtained during presentation of the reversed vocalization would be identical to the natural vocalization spike trains. In this case, application of the metric-space method to this new set of spike trains would result in nonsignificant amounts of transmitted information ( $H^*_{\text{max}} \sim 0$ ).

Reversed versions of reversed vocalization spike trains were constructed by considering the end of the trial (end of the vocalization) as the beginning of the trial. At this stage of the procedure, the end of the reversed vocalization was considered as the start of the new spike trains (time 0) and the last spike was considered as the first spike, the penultimate spike as the second spike, etc. The time shift, estimated on the basis of the latency of the first spike obtained in response to the natural vocalization, was added to each new spike time.

### Analysis of spike-timing reliability

The metric-space method can be used to analyze whether the spike train temporal structure significantly contributes to the transmitted informa-

tion and the particular values of temporal precision that maximize the transmitted information ( $q_{\max}$ ). It does not, however, reveal whether (1) a temporal organization of spike trains exists for the first, for the second, or for both stimuli (here the natural and time-reversed vocalizations) and (2) the reliability of the potential temporal organizations detected in the spike trains of both stimuli. To quantify the temporal reliability of neuronal responses on presentation of the natural and time-reversed vocalizations, a measure of correlation between spike trains was computed using a previously described method (Schreiber et al., 2003). Spike trains were first convolved with a Gaussian filter of a given width  $\sigma$ . The inner product was then computed between all pairs of trials and each inner product was then divided by the norms of the two trials of the respective pair. Reliability,  $R_{\text{corr}}$ , is the average across all pairs of trials of this “normed inner product.” Therefore, each convolved spike train  $s_i$  was represented as an individual vector and the correlation measure  $R_{\text{corr}}$  was given by the following:

$$R_{\text{corr}} = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i+1}^n \frac{\vec{s}_i \cdot \vec{s}_j}{\|\vec{s}_i\| \cdot \|\vec{s}_j\|},$$

where  $n$  is the number of stimulus presentations. The width of the Gaussian filter was set to 10 ms (according to the mean  $q_{\max}^*$  value that was derived from the metric-space analysis).

Trial-to-trial reliability of temporal patterns can differ between the four vocalizations simply because of their different durations. Indeed, phasic onset responses occurring in the first tens of milliseconds are probably more precise than later responses. Therefore,  $R_{\text{corr}}$  values simply computed for the entire stimulus duration would favor the shortest vocalizations. Here, to obtain  $R_{\text{corr}}$  values more comparable between stimuli, we computed  $R_{\text{corr}}$  values using a sliding window from the beginning to the end of each response. We chose a window of 80 ms, slid by 10 ms intervals, to have at least two values for the responses to the shortest vocalization, the chirp. This procedure led to several  $R_{\text{corr}}$  values (depending on the stimulus duration). From these values, we computed the average over all windows [ $R_{\text{corr,mean}}$ ] and the maximum [ $R_{\text{corr,max}}$ ]. For each stimulus, the  $R_{\text{corr,mean}}$  value, ranging from 0 (no reliability) to 1 (perfect reliability), quantifies the capability of a neuron to emit identical spike trains during successive presentations of the same stimulus over the entire duration of the stimulus. The  $R_{\text{corr,max}}$  quantifies the best spike-timing reliability that can be achieved by a neuron over the 80 ms sliding window.

All computations were made with Matlab and information theoretic analyses were conducted with the Spike-Train-Analysis-Toolkit available online (<http://neuroanalysis.org/toolkit/>) and statistical tests were made with Statistica (Statsoft). Statistical analyses involved ANOVAs (with “condition,” “type of vocalization,” and “direction of the vocalization” as factors) followed by *post hoc* comparisons. In all cases,  $p < 0.05$  was considered statistically significant.

### Histologic analyses

After each experiment performed in anesthetized animals, the brains were removed from the skull and placed in a fixative solution (4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4) for 2 weeks. For animals recorded in the awake state, the animals received a lethal dose of pentobarbital (200 mg/kg) after the final recording session, and small electrolytic lesions were made by passing anodal current (10  $\mu$ A, 10 s) through the recording electrodes. The animals were perfused transcardially with 0.9% saline (200 ml) followed by 2000 ml of fixative (4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4).

The brains were placed in a 30% sucrose solution in 0.1 M phosphate buffer for 3–4 d; then, coronal serial sections of the brain were cut on a freezing microtome (50  $\mu$ m thick). All serial sections were mounted on glass slides, dried, and counterstained with cresyl violet. The analysis of histologic material was performed blind to the electrophysiologic results. The sections were examined under several microscopic magnifications to find the electrode penetration tracks. For the MGB recordings, we located the recording sites using the electrode tracks for guidance, the point of entrance in the thalamus, the estimated dorsoventral extent of the MG,

and the depth coordinates read from the microdrive during the experiment. In all cases, this analysis was precise enough to allow clear assignment of the recordings with regard to the MG divisions. For the cortical recordings in anesthetized animals, the depth coordinates read from the microdrive and recent determinations of the relative thickness of cortical layers in the guinea-pig ACx (Wallace and Palmer, 2008) were used to assign each recording to a cortical layer. Both in pilot experiments and in previous studies (Manunta and Edeline, 1999; Edeline et al., 2001), there was good correspondence between the value read on the microdrive and the actual depth of the small electrolytic lesions made via the tungsten electrodes.

## Results

### Acoustic properties and localization of the recorded cells

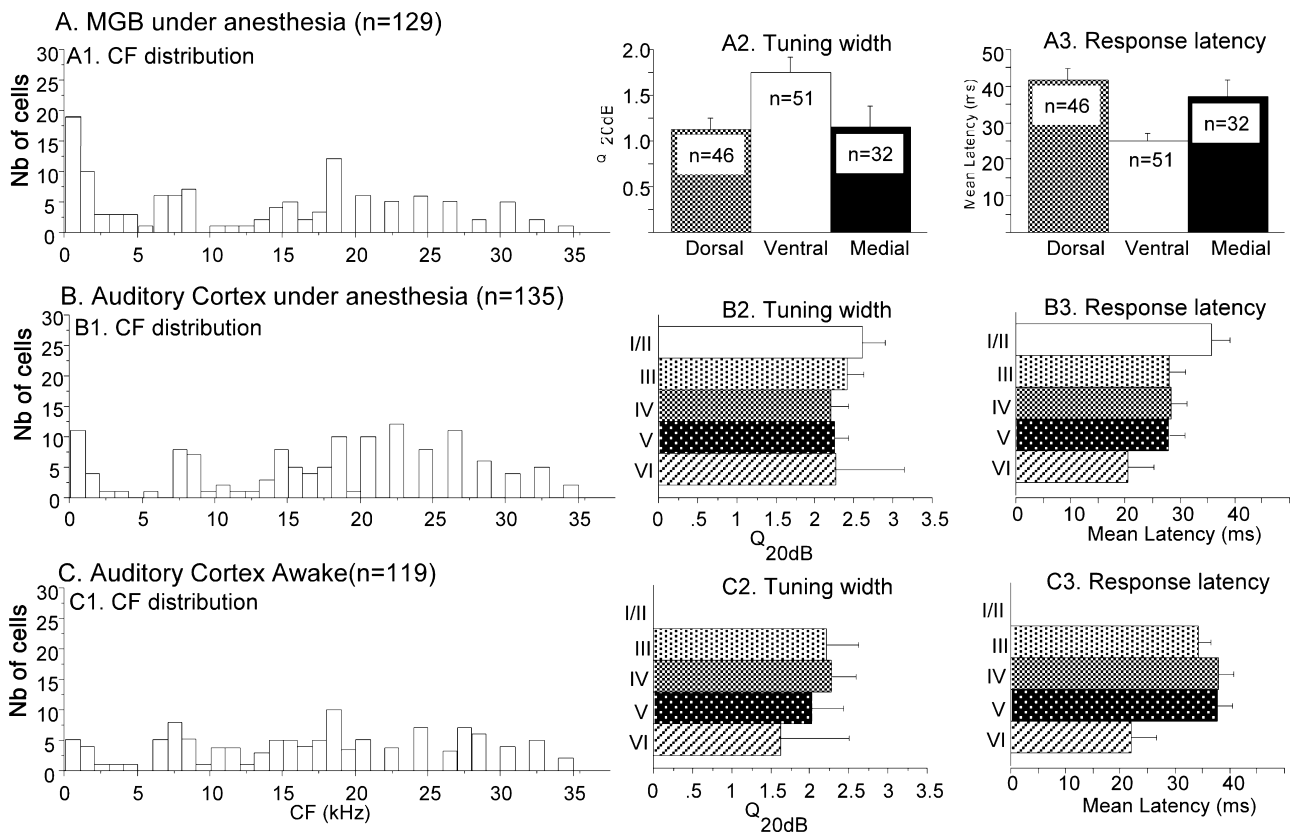
In anesthetized animals, stable single-unit recordings were obtained from 129 MGB cells and 135 cortical cells. For each animal, 5–18 cells were collected (mean 10 cells/animal). On average, spontaneous activity ranged from 0.01–5.3 spikes/s (median 0.58; mean  $1.08 \pm 0.46$ ); the responses at the best frequency ranged from 10.8–56.5 spikes/s (median 16.5; mean  $18.5 \pm 5.7$ ). Both in MGB and ACx, CF values ranged from 0.3–35 kHz (Fig. 2A1,B1). Fewer cortical cells were tuned to low frequencies ( $< 1$  kHz), but the two distributions did not significantly differ ( $\chi^2 = 1.25$ ; ns). Note that in primary ACx, cells with a CF  $< 1$  kHz are located rostrally (Wallace et al., 2000), in an area in which large blood vessels make long-lasting ( $> 40$  min) single-unit recordings particularly difficult to achieve.

In the MGB, the location of most of the cells relative to the MGB divisions was determined. On average, cells of the ventral MGB displayed sharper tuning curves (higher  $Q_{20\text{ dB}}$ ) and shorter response latencies (Fig. 2A2,A3). In the ACx, cells were recorded from 320 to 1850  $\mu$ m below the pia. Each cell was assigned to a cortical layer using the deep coordinates recently determined in the guinea pig ACx by Wallace and Palmer (2008). The sharpness of the tuning curves (Fig. 2B2) and the response latency (Fig. 2B3) did not significantly differ between cortical layers (ANOVA, lowest value  $p = 0.22$ ).

In awake animals, stable recordings from unambiguously isolated single units were obtained from 119 cortical cells (9–15 cells/animal, 3–4 recording sessions/animal, 3–5 cells/session). During each recording session, only one single unit was isolated from the signal obtained via a given electrode and only 1–3 electrodes (of 5–8 electrodes) produced signals allowing single-unit isolation. The CF of recorded cells ranged from 0.9–35 kHz (Fig. 2C1); the distribution did not significantly differ from that in the ACx of anesthetized animals ( $\chi^2 < 1$ ; ns). Spontaneous firing rate ranged from 0.1 to 11.6 spikes/s (median 1.32, mean  $2.25 \pm 0.25$ ) and was significantly higher than in anesthetized animals (unpaired *t* test,  $p < 0.01$ ). The strength of evoked responses at the best frequency ranged from 15.3–70.5 spikes/s (mean  $27.5 \pm 9.1$ ), which was not significantly different from that in the anesthetized animals (unpaired *t* test,  $p = 0.15$ ). Figure 2, C2 and C3, shows the laminar distribution of the sharpness of tuning and the response latency; there was no statistical difference between layers (ANOVA, lowest value  $p = 0.15$ ).

### Firing rate response to natural and time-reversed versions of the calls

Although all the cells included in the analyses displayed clear responses to pure tones, not all of them displayed reliable responses to vocalizations. Depending on the vocalization, the percentage of responsive cells (see Materials and Methods for definition) varied between 29 to 57%. Except for the purr, the percentages of responsive cells did not differ between MGB and



**Figure 2.** Characteristics of the recorded cells on presentation of pure tones in the auditory thalamus (A), the auditory cortex of anesthetized animals (B), and the auditory cortex of awake animals (C). The distribution of the characteristic frequency (CF) of cells did not differ between auditory thalamus (A1) and auditory cortex (B1) under anesthesia. In ACx, it did not differ between anesthetized (B1) and awake animals (C1). The tuning width quantified by the  $Q_{20\text{dB}}$  was higher in the tonotopic (MGv) division of the auditory thalamus than in the nontotopic (MGm and MGd) divisions (A2). They did not differ between cortical layers under anesthesia (B2) or in the awake condition (C2). The mean response latency was shorter in ventral auditory thalamus than in the dorsal and medial divisions (A3). It did not differ between cortical layers in anesthetized (B3) or awake (C3) condition. Under anesthesia, there was a tendency for longer latencies in the superficial layers (I/II) and shorter latencies in the deep layer (VI), but this effect was not significant.

**Table 1. Numbers and percentages of cells (1) that were considered "responsive" to each vocalization, (2) that carried information about each vocalization (vs its backward version), and (3) that carried information based on the temporal organization of the spike trains**

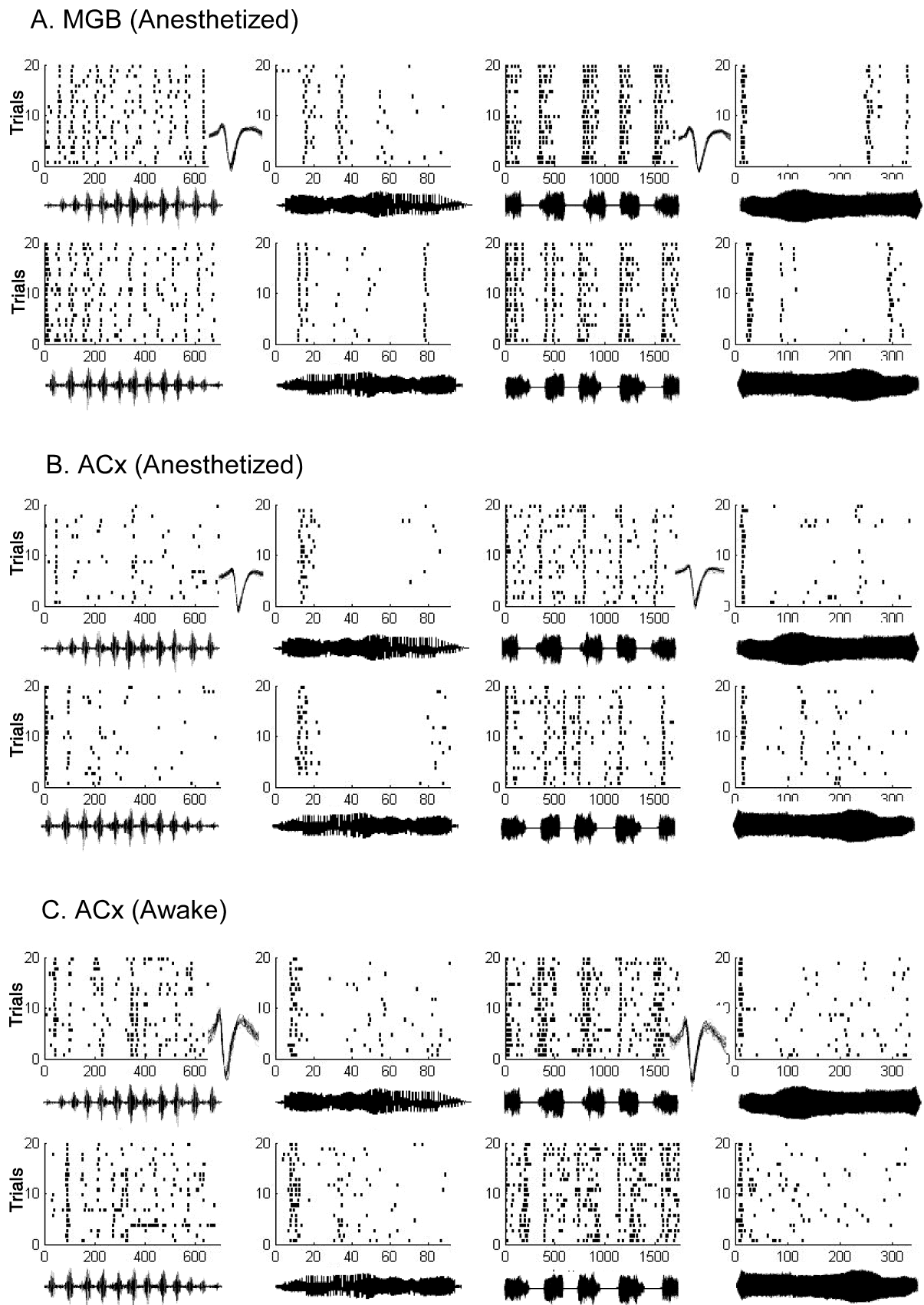
Stimulus (duration)	Condition	All cells <i>n</i>	Responsive cells		Responsive and informative cells		Responsive and informative and temporal cells	
			<i>n</i>	% of all cells	<i>n</i>	% of responsive cells	<i>n</i>	% of responsive and informative cells
Chutter (1740 ms)	MGB-anest	129	69	53	62	90	60	97
	ACx-anest	135	60	44	49	82	43	88
	ACx-awake	119	49	41	46	94 <sup>#</sup>	45	98 <sup>#</sup>
Purr (700 ms)	MGB-anest	129	60	46	53	88	50	94
	ACx-anest	135	39	29 <sup>**</sup>	36	92	32	89
	ACx-awake	119	39	33	38	97	38	100 <sup>#</sup>
Whistle (335 ms)	MGB-anest	129	68	53	50	73	38	76
	ACx-anest	135	77	57	55	71	45	82
	ACx-awake	119	66	56	29	44 <sup>#</sup>	28	97 <sup>#</sup>
Chirp (90 ms)	MGB-anest	129	61	47	35	57	27	77
	ACx-anest	135	76	56	32	42	24	75
	ACx-awake	119	66	56	27	41	26	96 <sup>#</sup>

Responsive cells were defined as cells exhibiting obvious responses in their rasters and poststimulus time histograms as assessed by two of the authors. Informative cells were defined as cells exhibiting positive values of  $H_{\text{max}} - (H_{\text{bias}} + 2 \text{ SD bias})$ . Temporal cells were defined as cells exhibiting positive values of  $H_{\text{max}}^*$  for a value of  $1/q_{\text{max}}$  inferior to the stimulus duration. \*\* indicates the condition for which the percentages of ACx cells significantly differed from the percentages of MGB cells in anesthetized animals ( $\chi^2, p < 0.05$ ). # indicates the conditions for which the percentages of ACx cells in awake animals significantly differed from the percentages of ACx cells in anesthetized animals ( $\chi^2, p < 0.05$ ; # $p < 0.06$ ).

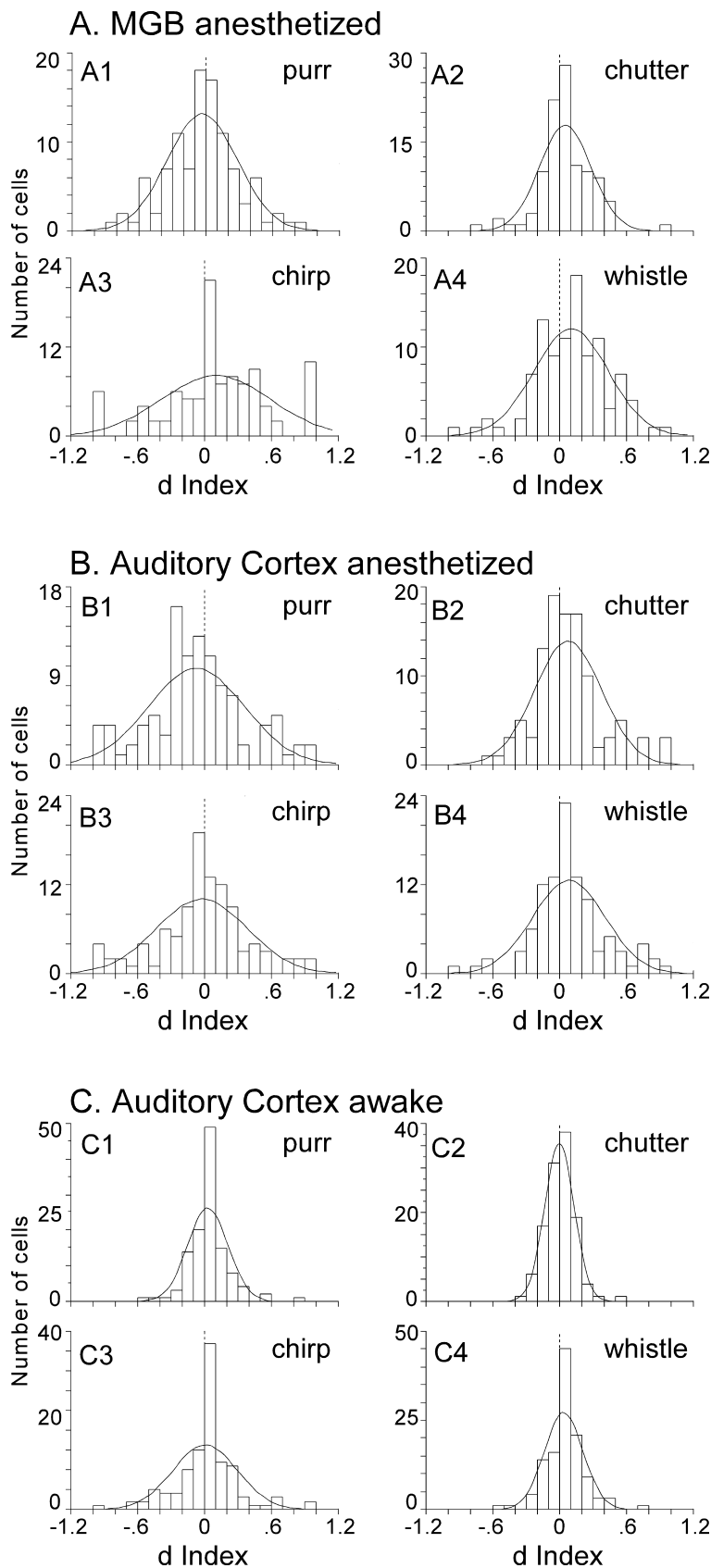
ACx (Table 1). The percentages of responsive cells obtained in the ACx of awake animals were similar to those obtained in anesthetized animals (all  $\chi^2$  values  $< 1$ , ns). The following paragraphs focus on responsive cells, but analyses of the cells classified as nonresponsive are presented below.

For most of the cells, when strong evoked responses were obtained on presentation of the natural vocalization, robust

responses were also obtained on presentation of the time-reversed vocalization. Figure 3 displays rasters obtained on presentation of the natural and time-reversed versions of the four vocalizations for one MGB cell (Fig. 3A), one ACx cell recorded under anesthesia (Fig. 3B), and one ACx cell recorded in the awake condition (Fig. 3C). These three cells responded to the four vocalizations and their responses were



**Figure 3.** Examples of single-unit responses obtained from anesthetized animals in the auditory thalamus (*A*) and auditory cortex (*B*), and from awake animals in auditory cortex (*C*). For each cell, the raster displays the responses to the natural version (top, 20 repetitions) and the time-reversed version (bottom, 20 repetitions) of the vocalizations. The oscillogram of the vocalizations is presented below each raster. Note that for each cell, the firing rates on presentation of the natural and time-reversed vocalizations were virtually identical. Note also that striking temporal patterns of discharge were observed on presentation of natural and time-reversed vocalizations. The waveform of the action potential (30 sweeps) obtained during the first (left) and last (right) trials is presented for each cell.



**Figure 4.** Distribution of the selectivity index ( $d$ ) for each vocalization and each condition. Each histogram shows the distribution of the selectivity index ( $d$ ) for a particular vocalization over the whole population of thalamic (**A**) or cortical (**B, C**) cells. A  $d$  value of 1 (–1) indicates that a neuron responded only to the natural (reversed) call; 0 means no firing-rate difference between natural and reversed calls. None of the distributions differed from a normal distribution centered around a mean of zero. Note that the values obtained for the cells recorded in awake animals were more concentrated around a mean of zero.

just as vigorous to the time-reversed versions as to the natural versions.

The distributions of the selectivity index  $d$  (see Materials and Methods) at the thalamic and cortical levels are presented in Figure 4 for each vocalization. None of these distributions differ from a normal distribution centered on zero (in all cases  $\chi^2 < 1$ ; ns). For each stimulus, “selective” cells exhibiting high positive or negative  $d$  values ( $d > 0.8$ ) were rare, and similar numbers of cells responded more vigorously to the natural or to the time-reversed version. The distributions obtained in MGB and ACx were not significantly different ( $\chi^2$ ; lowest value,  $p = 0.14$ ). In contrast, the distributions obtained in awake and anesthetized animals were significantly different ( $\chi^2$ , highest  $p$  value for the chirp,  $p = 0.03$ ): the distributions of  $d$  values were more concentrated around zero in awake animals, and none of the cells recorded in the awake conditions exhibited a high  $d$  value for either the natural or the time-reversed vocalizations.

**Information theory analyses**

The metric-space analysis was applied to spike trains of all recorded cells in the MGB and ACx. The responses of a given cell to a given vocalization were classified as “rate responses” or as “temporal responses” depending on the temporal precision ( $q^*_{max}$ ) value that provided the maximal amount of information ( $H^*_{max}$ ).

*Examples of rate responses and temporal responses*

Figure 5 presents examples of rate responses from four different cells. The rasters show that each cell responded only to either the natural (Fig. 5A1,B1) or the time-reversed (Fig. 5A2,B2) version of the stimulus. In the right column of Figure 5, the amount of transmitted information is plotted for a temporal precision  $q$  ranging from 0 to  $10^{-5} \text{ s}^{-1}$  in logarithmic steps. On such curves, the point at which  $H^*$  reaches its highest value ( $H^*_{max}$ ) indicates the extent to which a correct discrimination between stimuli can be achieved by looking at the spike trains, and the temporal precision with which this discrimination is achieved ( $q^*_{max}$ ). For all these curves,  $H^*$  reached its highest value (between 0.49 and 0.85 for these examples) for  $q^*_{max} = 0$ , indicating that information was maximal when only spike count was considered. For these cells, as  $q$  increases (and  $1/q$  decreased), the value of  $H^*$  gradually



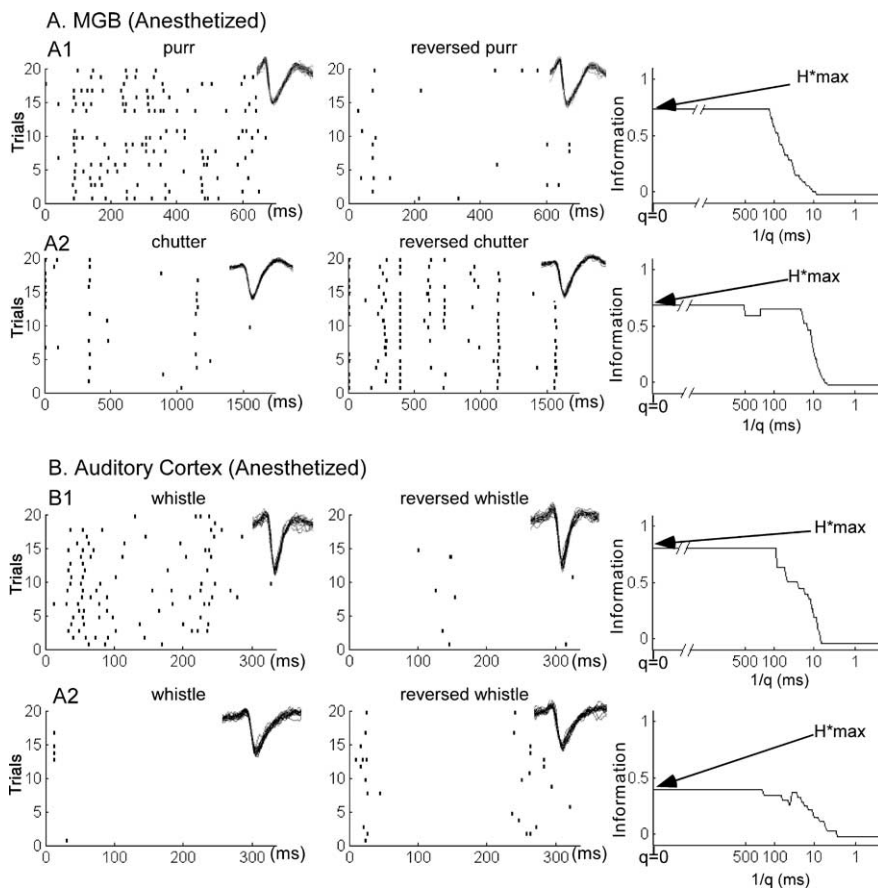
decreased, indicating that spike timing did not provide a better discrimination between natural and time-reversed vocalizations.

For most of the cells exhibiting significant values of  $H^*_{\max}$  (i.e., positive value of  $[H^*_{\max} - (H^*_{\text{bias}} + 2 \text{SD})]$ ), the responses were classified as “temporal responses.” Figure 6 presents six examples of temporal responses: temporal organization of evoked discharges emerged during an equivalent firing rate to natural and time-reversed stimuli. As shown in the last column of Figure 6, in all cases,  $H^*$  increased as  $q$  increases (and  $1/q$  decreased) up to a particular value of  $q$ , which indicated that the natural and reversed vocalizations produce different temporal patterns and that these differences were sufficiently reliable to discriminate between the two stimuli, whereas the firing rate was not.

#### Group data

A large majority of responsive cells (exhibiting clear evoked responses based on the rasters and post stimulus time histograms) were informative cells, i.e., cells for which the information value  $H^*_{\max}$  was significantly positive. Table 1 presents for each stimulus type, and each condition, the percentage of responsive cells, of informative cells, and of temporal cells, i.e., cells for which spike timing carries more information than spike count. Overall, the percentage of informative cells (cells that carried a significant amount of information for the four vocalizations) did not differ between MGB and ACx ( $\chi^2 < 1$ ; ns). A large majority of informative cells were temporal cells and the overall percentage of such cells did not differ between the MGB and ACx ( $\chi^2 = 2.13$ ;  $p = 0.14$ ). Among the responsive cells, the overall percentage of informative cells did not differ between the anesthetized and awake conditions in ACx ( $\chi^2 = 1.97$ ;  $p = 0.16$ ). In contrast, there was a globally higher proportion of temporal cells in the population of cells recorded in the awake condition compared with the anesthetized condition ( $\chi^2 = 6.56$ ;  $p = 0.01$ ). The effect was present for each vocalization (Table 1).

Figure 7 presents, for all responsive cells, the value of  $H^*_{\max}$  as a function of  $H^*_{\text{count}}$  for each vocalization and each condition. The dots located on the diagonal lines correspond to “rate responses,” i.e., cases for which  $H^*_{\max} = H^*_{\text{count}}$ . For each scattergram, most of the dots are located above the diagonal line ( $H^*_{\max} > H^*_{\text{count}}$ ), which means that taking the temporal precision into account increased the amount of transmitted information. In awake animals, the  $H^*_{\text{count}}$  values are quite small, whereas the  $H^*_{\max}$  values are similar to those obtained in anesthetized animals, which leads, for all stimuli, to clusters of dots around the vertical line corresponding to  $H^*_{\text{count}} = 0$ . Table 2 presents the mean values of  $H^*_{\text{count}}$  and  $H^*_{\max}$  for each stimulus, each structure, and each condition. Under anesthesia, the mean value of  $H^*_{\text{count}}$ ,  $H^*_{\max}$ , and  $q^*_{\max}$  obtained for the four vocalizations did not differ between MGB and ACx. In contrast, the mean  $H^*_{\text{count}}$  values obtained in the ACx of awake animals were lower than



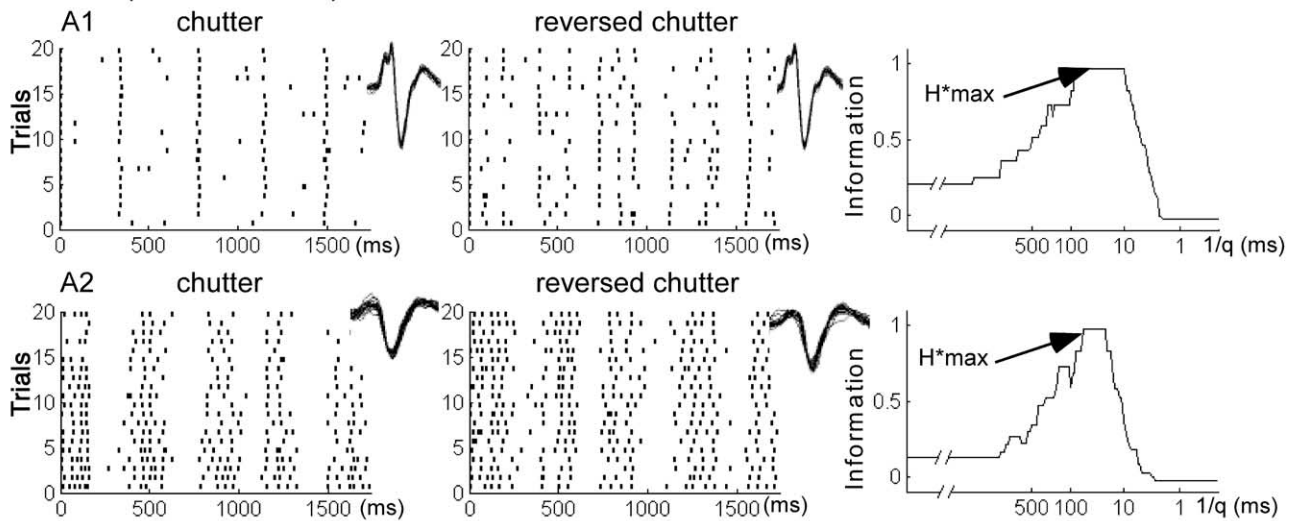
**Figure 5.** Raster plots and transmitted information for four cells exhibiting “rate responses.” For each cell, the raster shows the responses to 20 repetitions of the natural version (left column) and the time-reversed version (middle column) of a vocalization. The curves of the right column show the transmitted information as a function of the temporal precision ( $1/q$ ). In these four cases the highest value of information ( $H^*_{\max}$ ) is achieved when no temporal precision is required ( $q = 0$ ), and when  $q$  increases ( $1/q$  decreases) the information rapidly declines and reaches values close to zero. The waveform of the action potential (30 sweeps) obtained during the first (left) and last (right) trials is presented for each cell.

those obtained in anesthetized animals (ANOVA,  $p < 0.02$ ); in fact, information carried by spike count was almost null in awake animals. There was a significant interaction between type of stimulus and condition ( $p < 0.01$ ): For the two shorter stimuli (chirp and whistle), the mean  $H^*_{\max}$  values obtained in awake animals were lower than those obtained in anesthetized animals ( $p < 0.01$ ). The reverse was true for the two longer stimuli (purr and chutter): the mean  $H^*_{\max}$  value was higher in the awake than in the anesthetized animals. This suggests that, in awake animals, temporal discharge patterns can underlie stimuli discrimination if the stimuli are at least a few hundred milliseconds in duration.

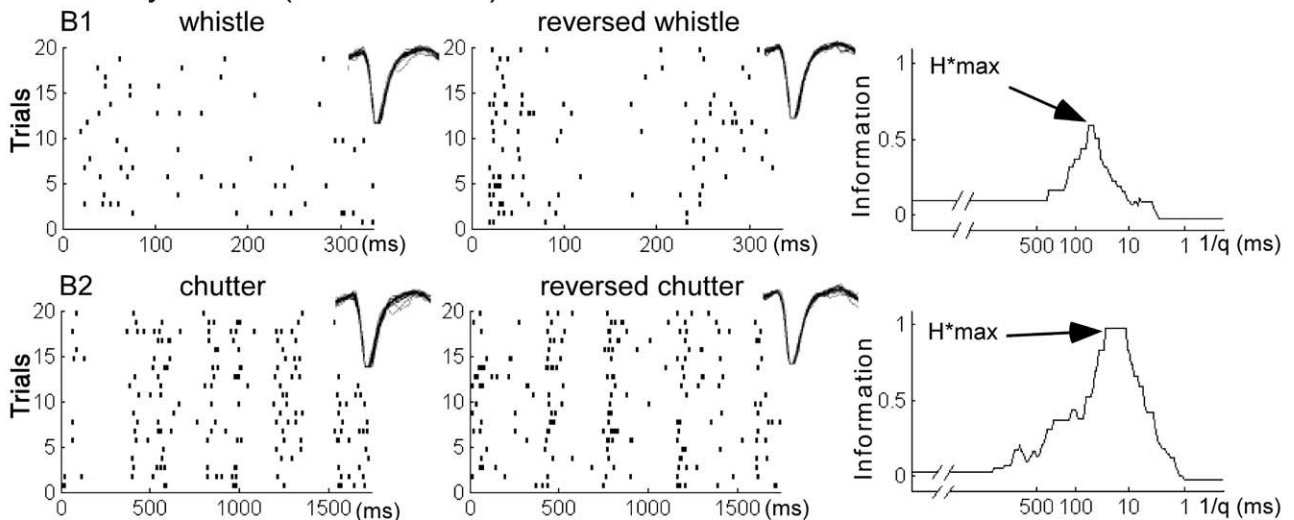
#### Analyses after reversing the spike trains of the reversed vocalizations

For a vast majority of the cells, visual examination of the rasters indicated that the temporal organization occurring on presentation of the time-reversed vocalization was not the mirror image of the one obtained with the natural vocalization (see examples in Figs. 3 B, C, 6 A–C). To substantiate this observation, we reversed the spike trains collected on presentation of the time-reversed stimuli and, after shifting these reversed spike trains by the response latency, we recomputed the information between these new sets of spike trains and those obtained from the natural stimuli. This analysis was performed for all cells that did not exhibit a pure onset response (for this reason, the responses to the chirp were discarded). For most of the cells (>75%),  $H^*_{\max}$  was

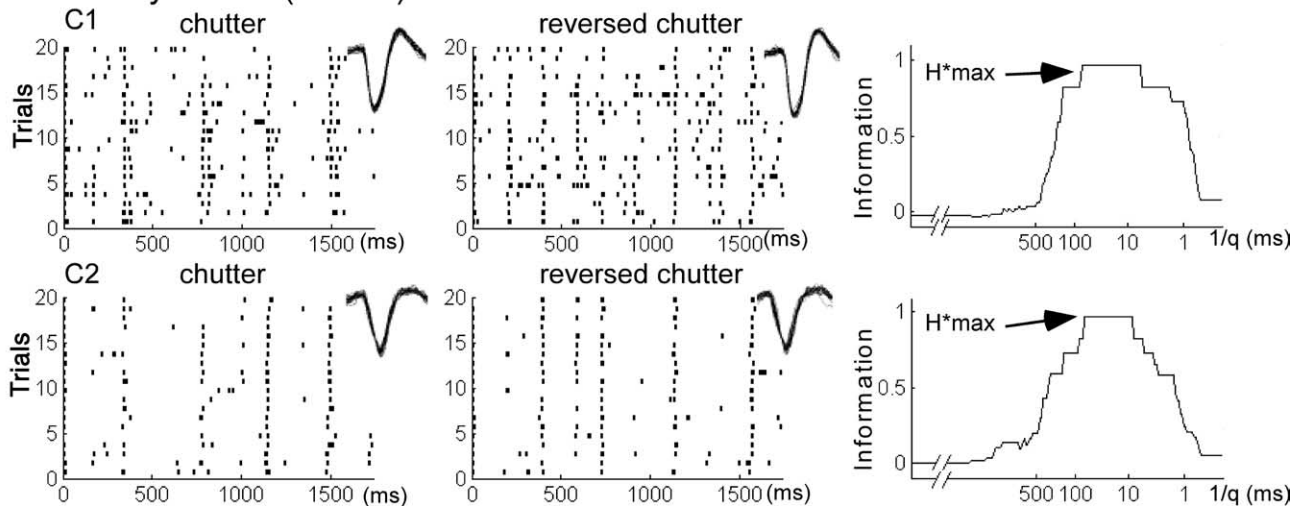
**A. MGB (Anesthetized)**



**B. Auditory Cortex (Anesthetized)**



**C. Auditory Cortex (Awake)**



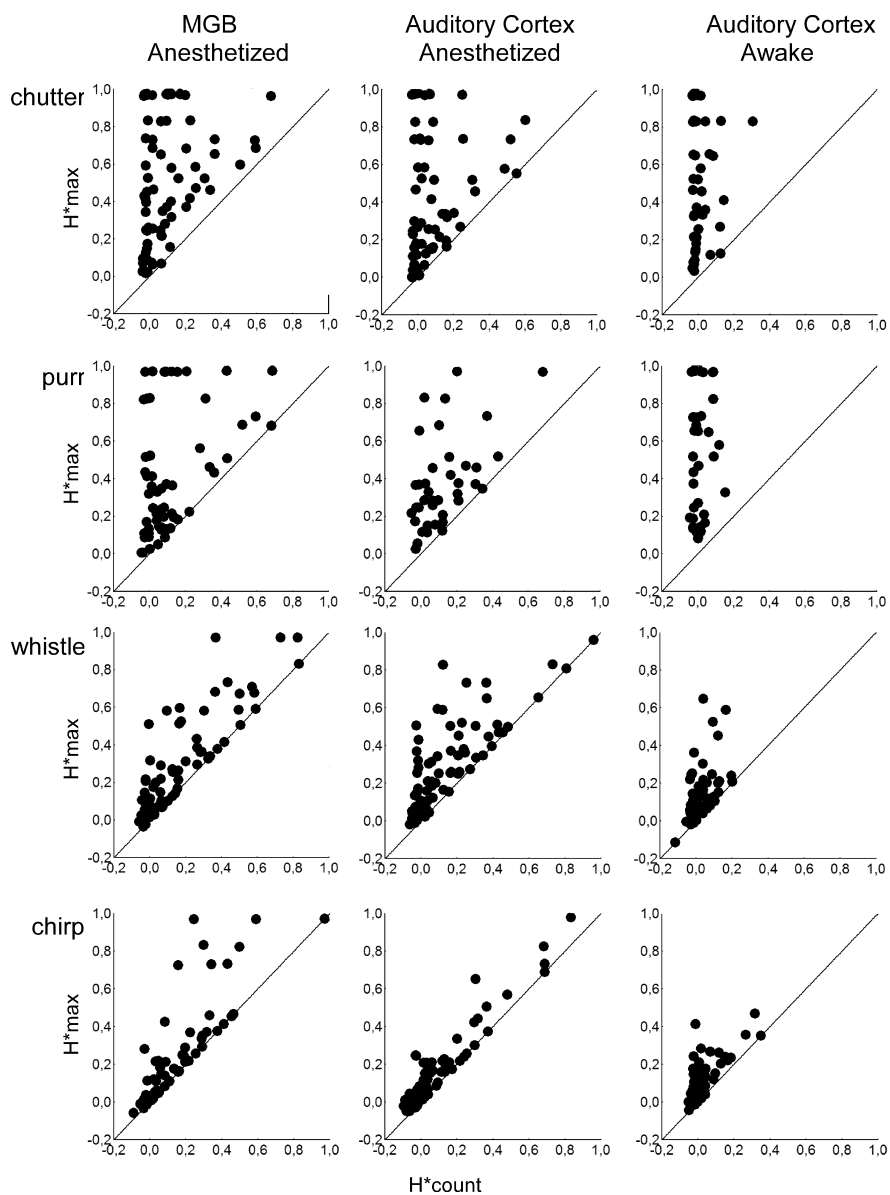
**Figure 6.** Raster plots and transmitted information for cells exhibiting “temporal responses.” For each cell, the raster shows the responses to 20 repetitions of the natural version (left column) and the time-reversed version (middle column) of a vocalization. The curve of the right column shows the transmitted information as a function of the temporal precision ( $1/q$ ). Note that the value of transmitted information is low when no temporal precision is considered ( $q = 0$ ) and that the highest value of information ( $H^*_{max}$ ) is achieved for a particular value of  $1/q$ ; then the information rapidly declines and reaches values close to zero. Note also that in all cases, the values of  $1/q$  corresponding to  $H^*_{max}$  are in the range of tens of milliseconds. The waveform of the action potential (30 sweeps) obtained during the first (left) and last (right) trials is presented for each cell.

not decreased but rather increased after this inversion (e.g., lowest  $p$  value = 0.042 for the purr). This means that reversing and shifting the spike train obtained from the time-reversed vocalization produced a spike train that differs more from the spike train obtained from the natural vocalization than does the original spike train obtained from the reversed vocalization.

### Spike-timing reliability of responses to vocalizations

The metric-space method indicated that high values of information were obtained when temporal discharge patterns were considered, but use of this method did not indicate whether temporal regularities existed for the natural stimulus, its time-reversed version, or both. Quantification of spike-timing reliability using the  $R_{\text{corr}}$  index allowed for the presence of temporal regularities to be assessed. For each vocalization, the  $R_{\text{corr,mean}}$  was similar for the responses to natural and to time-reversed stimuli (Fig. 8). An ANOVA indicated no effect of the factor “direction of the stimulus” ( $F_{(1,297)} = 1.24$ ;  $p = 0.27$ ). The chirp used in our study provided higher  $R_{\text{corr,mean}}$  values than the other vocalizations, but this was most likely because of the different length of the vocalizations. Except for the chirp, the vocalizations lasted long enough to compute the  $R_{\text{corr,mean}}$  value obtained over many periods of time, which attenuates the high spike-timing reliability obtained at some time points. For the chirp (the one used here as well as any chirps of similar duration), only two values were available, and as the chirp often produced very reliable onset responses, this led to high  $R_{\text{corr}}$  values that were not balanced by lower values when computing the  $R_{\text{corr,mean}}$  value. When we considered the maximal values,  $R_{\text{corr,max}}$ , obtained for each stimulus with a temporal sliding window, the values obtained for the chirp used here were not different (and somewhat lower) than the maximal values obtained for the other vocalizations (0.82 for the chirp, vs 0.87, 0.95, and 0.91 for the purr, chatter, and whistle, respectively). Note that when cells ( $n = 49$ ) recorded in awake conditions were tested with different samples of a particular vocalization type (e.g., the chatter), similar  $R_{\text{corr,mean}}$  values were obtained (see supplemental Fig. 1, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). Note also that choosing a value of 10 ms for the Gaussian filter,  $\sigma$ , allowed us to detect increases or decreases in the  $R_{\text{corr,mean}}$  value (see supplemental Fig. 2, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material).

The  $R_{\text{corr,mean}}$  values did not differ between cortex and MGB (Fig. 8). The  $R_{\text{corr,mean}}$  value was higher in the awake condition than in the anesthetized condition for the longer stimuli (purr



**Figure 7.** Scattergrams of the firing rate-based information versus the spike-timing-based information. For each stimulus and each structure and condition, the scattergrams represent the highest value of information when spike timing is considered ( $H^*_{\text{max}}$ ) as a function of the information obtained when only the firing rate is considered ( $H^*_{\text{count}}$ ). In each scattergram, the points on the diagonal line correspond to “rate responses” and the points above the diagonal line to “temporal responses.” For the four vocalizations and in the three conditions, a large number of cells displayed  $H^*_{\text{count}}$  values equal or close to zero, whereas the  $H^*_{\text{max}}$  values spread over the entire  $y$ -axis. In awake animals, these cells were so numerous that clusters of points generate a vertical line pattern.

and chatter). An ANOVA revealed a significant effect of the conditions ( $p < 0.01$ ) with no effect of the stimulus direction ( $p = 0.25$ ). *Post hoc* analyses revealed a significant difference between the anesthetized and awake conditions ( $p = 0.01$ ), but not between MGB and cortex under anesthesia ( $p = 0.35$ ).

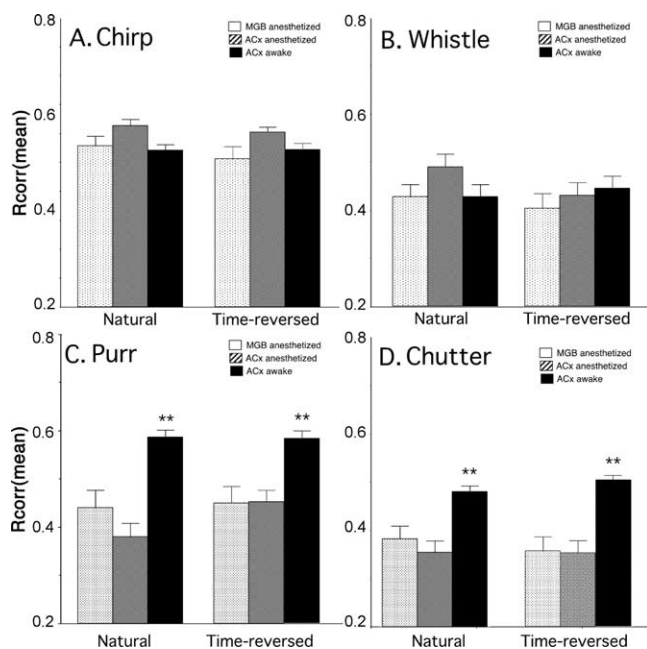
### Relations between $R_{\text{corr}}$ and $H^*_{\text{max}}$

On the one hand, the metric-space method indicated that for the four stimuli, the two structures (MGB and ACx) and the two conditions (anesthetized and awake), high values of information were obtained when spike timing was considered. However, the  $R_{\text{corr,mean}}$  values indicated that temporal organizations existed both on presentation of the natural and the time-reversed vocalizations. Although both of these measures depend on spike tim-

**Table 2.** Mean value of  $H_{\text{count}}$  and  $H_{\text{max}}$  for the informative cells and for the temporal informative cells

Stimulus (duration)	Condition	Responsive and informative cells		Responsive and informative and temporal cells	
		$H_{\text{count}}^*$	$H_{\text{max}}^*$	$H_{\text{max}}^*$	Precision 1/q (ms)
Chutter (1740 ms)	MGB-anest	0.11	0.57	0.58	71
	ACx-anest	0.11	0.47	0.48	118
	ACx-awake	0.01 <sup>##</sup>	0.56	0.57	32 <sup>#</sup>
Purr (700 ms)	MGB-anest	0.13	0.47	0.47	57
	ACx-anest	0.13	0.40	0.41	38
	ACx-awake	0.01 <sup>##</sup>	0.51	0.51	29
Whistle (335 ms)	MGB-anest	0.24	0.39	0.39	22
	ACx-anest	0.20	0.38	0.36	33
	ACx-awake	0.05 <sup>##</sup>	0.24 <sup>##</sup>	0.24 <sup>##</sup>	39
Chirp (90 ms)	MGB-anest	0.24	0.39	0.40	7
	ACx-anest	0.23	0.33	0.31	4
	ACx-awake	0.06 <sup>##</sup>	0.20 <sup>##</sup>	0.20 <sup>##</sup>	10 <sup>#</sup>

<sup>##</sup> indicates the cases for which the mean value obtained for ACx cells in awake animals significantly differed from the mean value obtained in anesthetized animals (unpaired *t* test,  $p < 0.01$ ; <sup>#</sup> $p < 0.05$ ).



**Figure 8.** Average  $R_{\text{corr,mean}}$  values obtained for each vocalization and each condition. The temporal reliability of neuronal responses at presentation of the natural and time-reversed vocalization was quantified by the  $R_{\text{corr}}$  index (Schreiber et al., 2003) (see Materials and Methods). For each vocalization and each condition, the  $R_{\text{corr,mean}}$  was computed with a temporal window of 80 ms shifted by 10 ms. There was no difference between cortex and thalamus in anesthetized conditions regardless of the vocalization. For the two longer stimuli (purr and chutter), the  $R_{\text{corr,mean}}$  values obtained in the awake conditions were higher than those obtained under anesthesia, but this effect was not observed for the two shorter stimuli (chirp and whistle).

ing, they do not quantify the same aspects of spike train organization. Indeed, the metric-space method relies on a comparison between spike trains elicited by different stimuli whereas  $R_{\text{corr}}$  quantifies the spike-timing reliability for the responses to each stimulus. A positive correlation between these two quantifications would therefore indicate that the temporal organization of spike trains emerges during presentation of both stimuli and the higher their temporal reliability ( $R_{\text{corr}}$ ), the better their ability to discriminate the stimuli ( $H_{\text{max}}^*$ ).

The scattergrams presented in Figure 9 show that for the two longer stimuli (purr and chutter), there was a correlation between the  $H_{\text{max}}^*$  and  $R_{\text{corr,mean}}$  values. Except for the values ob-

tained in the ACx under anesthesia on the presentation of purr ( $r = 0.14$ , ns), these relations were significant ( $p < 0.05$ ). There was no correlation for the chirp (highest value  $r = 0.11$ ; ns), and for the whistle there was an overall correlation that was not significant in all conditions (it was only significant in ACx under anesthesia). The correlations indicate that, provided the stimulus is long enough, the higher the spike-timing reliability the higher the amount of information carried by temporal discharge patterns. Temporal regularities occurring too briefly do not necessarily provide high information values.

### Relationship between anatomic location of the recorded cells and quantifications based on the information theory method

At the thalamic level, the different MGB divisions showed similar proportions of cells responding to the four vocalizations, suggesting that lemniscal and nonlemniscal MGB had similar proportions of cells responding to conspecific vocalizations (see also Philibert et al., 2005). The percentage of dorsal MG (MGd) cells displaying a significant amount of information was systematically lower than in the medial MG (MGm) and in the ventral MG (MGv). For example, this effect was significant for the purr ( $\chi^2$ ,  $p = 0.04$  for MGd vs MGm and MGd vs MGv). The effect was also significant when considering cells that exhibited significant amounts of information for all vocalizations (2/27 cells in MGd vs 12/39 cells in MGv and 7/20 cells in MGm;  $\chi^2$ ,  $p < 0.02$  for MGd vs MGv and  $p < 0.04$  for MGd vs MGm).

At the cortical level, no statistical differences were found for the between-layer comparisons. The proportions of cells responsive to the four vocalizations did not differ between the different layers. For all types of vocalization, the percentage of “informative” cells was comparable in each layer ( $\chi^2$ ; lowest  $p$  value = 0.09). Last, the percentage of “temporal” cells was higher in layer V than in the superficial layers for the whistle ( $\chi^2$ ;  $p = 0.02$ ) and for the purr ( $\chi^2$ ;  $p = 0.06$ ), but there was no significant effect for the other two vocalizations.

### The spike trains of nonresponsive cells can carry information

Surprisingly, the metric-space analysis revealed that a non-negligible percentage of cells classified as nonresponsive (exhibiting evoked discharges in  $< 10$  of 20 trials) had significantly positive  $H_{\text{max}}^*$  values. This finding suggests that, despite their poor firing rate, these cells displayed some temporal regularities in their spike trains that differed for the natural and time-reversed vocalizations. Figure 10A–C shows the rasters of three of these cells. These temporal regularities are less prominent than those observed in responsive cells, but they are sufficient to generate significantly positive  $H_{\text{max}}^*$  values. Figure 10D presents the mean  $H_{\text{max}}^*$  and  $H_{\text{count}}^*$  values for the responsive and nonresponsive cells. Whatever the vocalization was, the  $H_{\text{max}}^*$  values for the nonresponsive cells were as high, or even higher, than the  $H_{\text{count}}^*$  values for the responsive cells. This result suggests that selecting cells based only on the strength of the evoked discharges probably provides an underestimation of the number of cells whose spike trains contained a significant amount of information about the stimuli.

### Discussion

The vast majority of MGB and ACx cells recorded in anesthetized and awake guinea pigs displayed a similar evoked firing rate to natural and time-reversed conspecific vocalizations. Only a few cells produced spike trains whose firing rate carried a significant amount of information. However, for a large proportion of cells, the temporal organization of neural discharges carried more in-

formation than did spike count. Quantification of the spike-timing reliability revealed that there was no difference between natural and time-reversed vocalizations and no difference between MGB and ACx cells, but there were differences between awake and anesthetized conditions: for the two longer vocalizations, spike-timing reliability was higher in awake animals than in anesthetized animals.

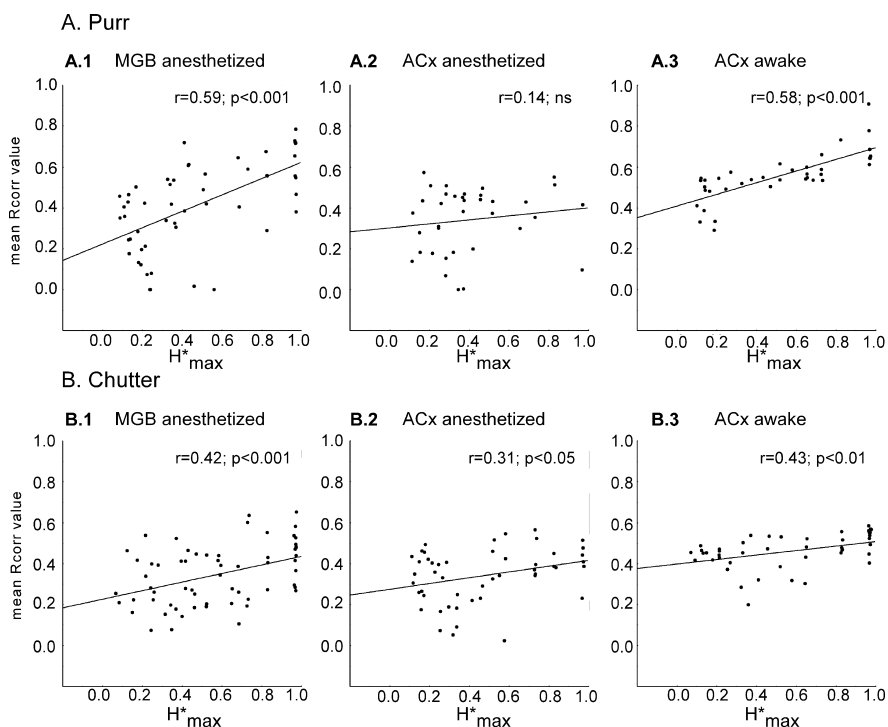
### Limitations of the present study

The metric-space analysis only applies to spike trains obtained from stimuli of equal duration. Thus, we could not determine whether the firing rate or temporal organization of neuronal discharges differed between the four different vocalizations used here. Nevertheless, comparison of the firing rate and spike patterns obtained on presentation of vocalizations with such different durations (from 90 to 1740 ms) and spectral content is almost trivial: any cell in the central auditory system can provide the neuronal basis for discriminating between these vocalizations based on their duration and spectral content. Discriminating between stimuli with similar durations and similar energies in the frequency bands (i.e., the natural and time reversed stimuli used here) is a more difficult task for which spike timing provides a more reliable neural basis than firing rate.

In our study, single-unit recordings were not collected in secondary auditory areas. Using natural and time-reversed conspecific vocalizations in cat, Gourévitch and Eggermont (2007) observed differences between primary and nonprimary areas (posterior ectosylvian gyrus): neurons exhibiting sustained responses in AI showed no preference for natural calls whereas neurons in posterior ectosylvian gyrus did. Because our recordings were collected only in tonotopic areas, the existence of a potential gradient from a temporal to a rate code in the hierarchy of auditory cortical areas remains to be investigated. Note, however, that in the auditory thalamus no differences were detected between tonotopic and nontotopic areas in terms of both number of cells bearing information based on temporal patterns and spike-timing reliability quantified by the  $R_{\text{corr}}$  index.

### Comparison with studies quantifying firing rate during conspecific vocalizations

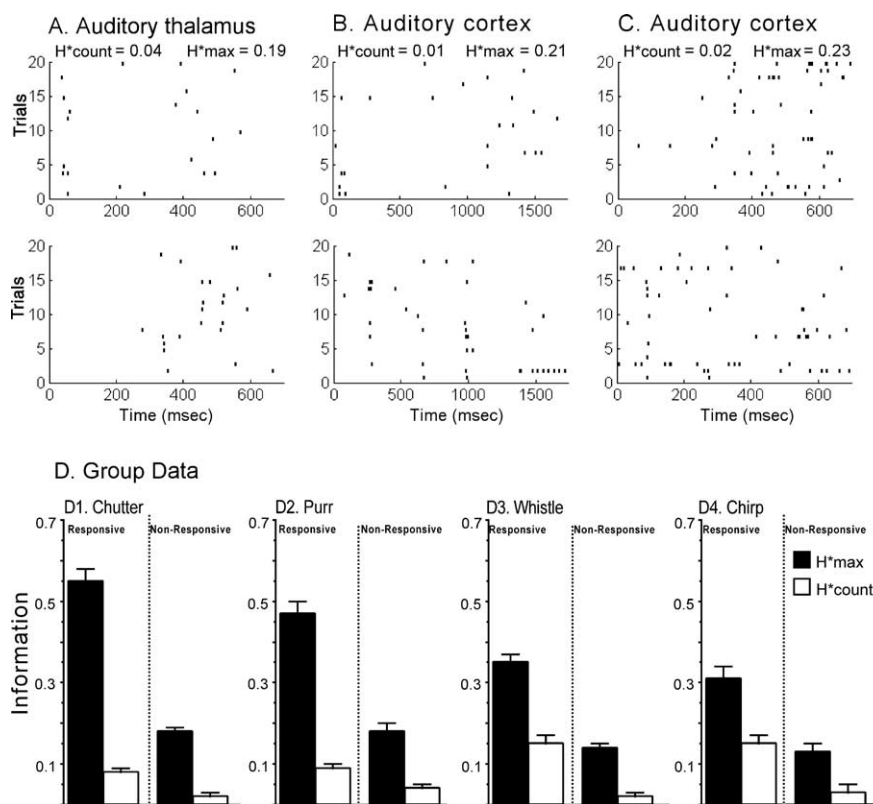
In songbird, intensive research has long focused on the neuronal responses to the BOS in different brain areas (review in Margoliash, 1997; Doupe and Kuhl, 1999; Prather and Mooney, 2004; Nealen and Schmidt, 2002). Selectivity for the BOS is weak in the avian homolog of primary auditory cortex (field L), but is robust in the HVC and Nif nuclei (Margoliash, 1986; Lewicki and Arthur, 1996; Janata and Margoliash, 1999; Cardin and Schmidt, 2004; Coleman and Mooney, 2004). In field L (and also in the caudal mesopallium), however, neurons respond preferentially to conspecific songs over synthetic sound ensembles designed to mimic the power spectra and amplitude modulation spectra of natural songs (see also Langner et al., 1981; Grace et al., 2003).



**Figure 9.** Relationship between the  $R_{\text{corr,mean}}$  value and the value of transmitted information based on the spike-timing  $H^*_{\text{max}}$ . For the two longer vocalizations (purr = 700 ms; chutter = 1740 ms), the scattergrams represent the  $R_{\text{corr}}$  values as a function of the  $H^*_{\text{max}}$  values. Except for the data obtained in auditory cortex in anesthetized animals, there was a relationship between the two variables: for a given cell, the higher the  $R_{\text{corr}}$ , the higher the  $H^*_{\text{max}}$ .

These data suggest that, in songbirds, ACx neurons respond preferentially to conspecific sounds.

In mammals, initial studies performed in the monkey ACx indicated that, based on firing rate, most cells (i) were not selective for a particular vocalization and (ii) responded similarly to the natural and time-reversed versions of vocalizations (Wollberg and Newman, 1972; Glass and Wollberg, 1983a,b; Pelleg-Toiba and Wollberg, 1991). In contrast, results obtained by Wang and colleagues in marmoset suggested that ACx neurons exhibit a firing rate preference for the natural version of conspecific calls (Wang et al., 1995), whereas when the same stimuli are presented to another species, the ACx neurons do not show this preference (Wang and Kadia, 2001). Our results are consistent with those obtained in the cat primary ACx (Gehr et al., 2000; Gourévitch and Eggermont, 2007), indicating that there is no firing rate preference for the natural version of conspecific stimuli. In the present study, very few thalamic and cortical cells (1–10%) showed a preference for the natural versions of the conspecific calls, and an equal proportion of cells showed a preference for the time-reversed versions. One might have assumed that testing cells under general anesthesia would interfere with the detection of a preference for the natural (behaviorally meaningful) version of the vocalizations, but, interestingly, the present results argued against this possibility. First, in terms of firing rate, neurons recorded in awake animals systematically showed similar responses to natural and time-reversed vocalizations (Fig. 4). Second, in terms of information content, there were more temporal responses, i.e., responses for which information is higher when spike timing is considered, in awake animals than in anesthetized animals (Table 1). Thus, in the thalamocortical auditory system of the awake animal, the neuronal bases



**Figure 10.** Individual examples and group data for the nonresponsive cells. **A–C**, Raster plots showing the lack of reliable responses for cells classified as “nonresponsive.” In each case, the raster shows that despite the fact that the cell was responding in  $<10$  of 20 of trials, when responses were emitted, the action potentials occurred during particular temporal windows that differed between natural and time-reversed vocalizations. **D**, Group data showing, for each vocalization, the mean value of transmitted information for the informative and temporal cells classified either as “responsive” (left two columns) or as “nonresponsive” (right two columns). The black bars represent the  $H^*_{\text{max}}$  value (when spike timing is considered), the open bars represent the  $H^*_{\text{count}}$  values (when only the firing rate is considered). Note that the  $H^*_{\text{max}}$  values obtained for the “nonresponsive” cells were as high as (or higher than) the  $H^*_{\text{count}}$  values obtained for the responsive cells.

for discriminating between natural and time-reversed vocalizations should rely on temporal discharge patterns.

### Comparison with studies evaluating the role of spike timing

In the visual modality, the temporal precision of neural discharges at the thalamic and cortical levels has been evaluated by several teams. A high temporal resolution (in the millisecond range) was initially found at the thalamic level on the presentation of sequences of image frames with a fluctuating luminance (Reinagel and Reid, 2000, 2002), but primary visual cortex (V1) neurons also fired with high precision on the presentation of visual stimuli (Kara et al., 2002; Kumbhani et al., 2007) (for review, see Tiesinga et al., 2008). Use of the metric-space method showed that the temporal organization of spike trains allowed for better discrimination of visual stimuli than firing rate, both in V1 (Victor and Purpura, 1996; Mechler et al., 1998) and V2 (Victor and Purpura, 1996). Recent studies also demonstrated that, during presentation of natural visual stimuli, significant amounts of information are carried by spike timing with millisecond and submillisecond precision (Nemenman et al., 2008). The gradual disappearance of spike-timing precision during progression through the sensory system is a plausible concept, but reanalysis of spike trains from MT neurons with new techniques has revealed that these neurons display quite high temporal precision in their firing patterns (Fellous et al., 2004).

In field L, mutual information computed on presentation of

complex acoustic stimuli indicates that a large amount of information is carried by spike timing, with approximately one-half of the information accessible only at time resolutions of  $\sim 10$  ms (Wright et al., 2002), a timescale that was optimal for neural discrimination of a conspecific song (Narayan et al., 2006). More recently, the evaluation of different candidate neural codes underlying song discrimination revealed that performance based on spike timing was higher than performance based on firing rate or interspike intervals (Wang et al., 2007). Last, in the bird telencephalic nucleus HVC, only 20% of the cells were classified as selective for the BOS based on firing rate, whereas among all the cells carrying information about the BOS and its time-reversed version, a large proportion (77%) could discriminate between the two versions based on the temporal information contained in the spike trains (Huetz et al., 2006).

In the mammalian ACx, initial studies using artificial stimuli indicated that spike patterns contained significantly more information about stimulus location than firing rate, both in anesthetized (Middlebrooks et al., 1994, 1998; Furukawa and Middlebrooks, 2002) and awake animals (Mickey and Middlebrooks, 2003). Recently, however, several studies have stressed the importance of temporal patterns of ACx neurons for coding communication sounds. For example, on presentation of heterospecific vocalizations, the highest amount of information carried by

spike patterns of ACx neurons was achieved at temporal resolutions between 10 and 20 ms and then the information content declined for temporal resolutions of 80 ms or greater (Schnupp et al., 2006). Similarly, information content about the identity of natural sounds was higher when quantified with temporal features of neuronal responses (such as the response latency, or binary patterns at a 4–8 ms resolution) than when quantified with spike count (Chechik et al., 2006). Consistent with the present study, Chechik et al. reported no difference in the average amount of information carried by spike trains of thalamic and cortical cells. In contrast, inferior colliculus neurons conveyed 2 to fourfold more information about the stimulus identity than ACx and MGB neurons, but they were more redundant than ACx and MGB neurons in the way they coded the stimulus identity (Chechik et al., 2006).

### What are the acoustic features and potential mechanisms that produce reliable spike timing?

If we assume that neurons are simple linear bandpass filters, action potentials should be emitted each time the vocalizations contain energy in a particular frequency band (but note that spectrotemporal receptive fields of ACx neurons can be quite sluggish, i.e., the temporal component of the spectrotemporal response field can be quite slow). The response nonlinearities, however, might be more prominent in higher auditory centers and some results suggest that ACx neurons respond to the higher

order statistics of a stimulus (Machens et al., 2004; Ahrens et al., 2008). Here, the analyses performed after inverting the spike trains obtained for time-reversed vocalizations allowed us to reject the possibility that the temporal discharge patterns were simply the consequences of a phase-locking on the stimulus envelope. Actually, data obtained by Heil and Irvine (Heil, 1997a,b; Heil and Irvine, 1996, 1998) indicate that rather than the envelope, it is the maximum acceleration of the peak pressure (i.e., the second derivative of a tone's envelope) that is responsible for the latency and the jitter of the first-spike response. In addition, ACx neurons can synchronize their discharges on the fine-structure of complex acoustic stimuli, responding not only to sound onset but also to the stimulus fine-structure with a precision of a few milliseconds (Elhilali et al., 2004, 2005). Therefore, the spike-timing precision observed on presentation of conspecific vocalizations can reflect a general mechanism for tracking rapid changes in peak pressure contained in most natural stimuli. This might explain why auditory thalamocortical neurons respond in a highly nonlinear manner, producing very different responses to small artificial modifications of a natural sound (Bar-Yosef et al., 2002; for review, see Nelken, 2004). Indeed, a large amount of information in the natural acoustic environment is carried by energy transients and this is particularly true for speech (Drullman, 1995; Shannon et al., 1995) and animal communication sounds. Hence, even if auditory cortex neurons do not process conspecific vocalizations differently from other natural sounds, they do process these vocalizations with a high degree of fidelity, thus allowing transients to be encoded.

Whichever acoustic features are responsible for the temporal discharge patterns, they must rely on precise and reliable mechanisms. Several mechanisms can contribute to the high spike-timing reliability of thalamocortical circuits (for review, see Tiesinga et al., 2008). First, *in vitro* studies indicate that cortical neurons display temporal regularities during injection of fluctuating currents in the soma (Mainen and Sejnowski, 1995; Fellous et al., 2001), suggesting that intrinsic membrane properties support high temporal fidelity. Second, recent studies suggest that auditory thalamocortical transmission is also temporally precise: the monosynaptic EPSPs triggered by thalamic stimulation in regular spiking and fast spiking cells from layers 3/4 can follow at rates of up to 40 Hz (Rose and Metherate, 2005). Third, the precise spiking pattern of ACx cells is often viewed as resulting from the precise temporal integration of excitatory and inhibitory potentials. Indeed, if inhibitory inputs are slightly delayed in time, ACx cells can emit action potentials during very brief temporal windows (Wehr and Zador, 2003; Zhang et al., 2003). Regardless of the contribution of these different mechanisms, they are not specific to cortical circuits because similar spike-timing precision is observed at the thalamic level. This is not surprising given that thalamic interneurons (Usrey and Reid, 1999) and thalamic reticular neurons (Cotillon-Williams et al., 2008) sculpt the temporal response of thalamic relay cells as cortical interneurons do in sensory cortices.

#### Importance of spike-timing precision for significant stimuli

The central question addressed by this study was how neural spike trains can discriminate a natural stimulus with a strong behavioral meaning for the animal's daily social interactions from a stimulus with a similar spectral and temporal content, but no behavioral relevance. Our results clearly indicate that temporal discharge patterns can provide the cellular basis for such a discrimination, whereas the firing rate cannot.

Two previous studies in mammals indicated that spike timing

is more efficient than firing rate for coding the behavioral relevance of vocalizations. For example, compared with responses recorded in the ACx of pup-naive females, the neuronal responses to pup calls obtained from mothers were not stronger in terms of firing rate but, at a 2 ms resolution, they contained more information about the detection and discrimination of pup calls (Liu and Schreiner, 2007). This effect was not observed for non-natural sounds falling outside the mouse vocalization repertoire. In another study, compared with the neuronal responses recorded from naive ferrets, the responses recorded in the ACx in ferrets that received discrimination training in a Go/NoGo task were not stronger for the conditioned stimulus, but they contained more information based on temporal discharge patterns expressed at timescales of 10–50 ms (Schnupp et al., 2006).

Obviously, reproducible spike-patterns observed at the single cell level cannot by themselves represent the neuronal basis of auditory perception, but their synchronization in neuronal ensembles is certainly a key element for the intercolumnar and interstructural communication (for review, see Tiesinga et al., 2008). The similar temporal pattern precision during natural and time-reversed vocalizations in the present study also argues against the existence of neurons that selectively express temporal organization for natural vocalizations that the animals are exposed to daily. Rather, our data indicate that different temporal discharge patterns are triggered by natural and artificial stimuli with the same spectral content. These temporal patterns are probably the reflection of transient neuronal assemblies that provide the neuronal bases of auditory perception.

#### References

- Ahrens MB, Linden JF, Sahani M (2008) Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectro-temporal methods. *J Neurosci* 28:1929–1942.
- Bar-Yosef O, Rotman Y, Nelken I (2002) Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *J Neurosci* 22:8619–8632.
- Berryman JC (1976) Guinea-pig vocalizations: their structure, causation and function. *Z Tierpsychol* 41:80–106.
- Calford MB, Webster WR, Semple MM (1983) Measurement of frequency selectivity of single neurons in the central auditory pathway. *Hear Res* 11:395–401.
- Cardin JA, Schmidt MF (2003) Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. *J Neurophysiol* 90:2884–2899.
- Cardin JA, Schmidt MF (2004) Auditory responses in multiple sensorimotor song system nuclei are co-modulated by behavioral state. *J Neurophysiol* 91:2148–2163.
- Chang EF, Merzenich MM (2003) Environmental noise retards auditory cortical development. *Science* 300:498–502.
- Chang EF, Bao S, Imaizumi K, Schreiner CE, Merzenich MM (2005) Development of spectral and temporal response selectivity in the auditory cortex. *Proc Natl Acad Sci U S A* 102:16460–16465.
- 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.
- Coleman MJ, Mooney R (2004) Synaptic transformations underlying highly selective auditory representations of learned birdsong. *J Neurosci* 24:7251–7265.
- Cotillon-Williams N, Edeline JM (2003) Evoked oscillations in the thalamo-cortical auditory system are present in anesthetized but not in unanesthetized rats. *J Neurophysiol* 89:1968–1984.
- Cotillon-Williams N, Edeline JM (2004) Evoked oscillations in unit recordings from the thalamo-cortical auditory system: an aspect of temporal processing or the reflection of hyperpolarized brain states? *Acta Neurobiol Exp (Wars)* 64:253–270.
- Cotillon-Williams N, Huetz C, Hennevin E, Edeline JM (2008) Topo-

- graphic control of auditory thalamus frequency tuning by reticular thalamic neurons. *J Neurophysiol* 99:1137–1151.
- de Villers-Sidani E, Chang EF, Bao S, Merzenich MM (2007) Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *J Neurosci* 27:180–189.
- Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* 22:567–631.
- Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am* 97:585–592.
- Edeline JM, Pham P, Weinberger NM (1993) Rapid development of learning-induced receptive field plasticity in the auditory cortex. *Behav Neurosci* 107:539–551.
- Edeline JM, Manunta Y, Hennevin E (2000) Auditory thalamus neurons during sleep: changes in frequency selectivity, threshold and receptive field size. *J Neurophysiol* 84:934–952.
- Edeline JM, Dutrieux G, Manunta Y, Hennevin E (2001) Diversity of receptive field changes in auditory cortex during natural sleep. *Eur J Neurosci* 14:1865–1880.
- Elhilali M, Fritz JB, Klein DJ, Simon JZ, Shamma SA (2004) Dynamics of precise spike timing in primary auditory cortex. *J Neurosci* 24:1159–1172.
- Elhilali M, Klein DJ, Fritz JB, Simon JZ, Shamma SA (2005) The enigma of cortical responses: slow yet precise. In: *Auditory signal processing: physiology, psychoacoustics, and models* (Pressnitzer D, de Cheveigné A, McAdams S, Collet L, eds), pp 485–494. New York: Springer.
- Engineer ND, Percaccio CR, Pandya PK, Moucha R, Rathbun DL, Kilgard MP (2004) Environmental enrichment improves response strength, threshold, selectivity, and latency of auditory cortex neurons. *J Neurophysiol* 92:73–82.
- Evans EF (1979) Neuroleptanesthesia for the guinea pig. *Arch Otolaryngol* 105:185–186.
- Fellous JM, Houweling AR, Modi RH, Rao RP, Tiesinga PH, Sejnowski TJ (2001) Frequency dependence of spike timing reliability in cortical pyramidal cells and interneurons. *J Neurophysiol* 85:1782–1787.
- Fellous JM, Tiesinga PH, Thomas PJ, Sejnowski TJ (2004) Discovering spike patterns in neuronal responses. *J Neurosci* 24:2989–3001.
- Furukawa S, Middlebrooks JC (2002) Cortical representation of auditory space: information-bearing features of spike patterns. *J Neurophysiol* 87:1749–1762.
- Gehr DD, Komiya H, Eggermont JJ (2000) Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. *Hear Res* 150:27–42.
- Glass I, Wollberg Z (1983a) Auditory cortex responses to sequences of normal and reversed squirrel monkey vocalizations. *Brain Behav Evol* 22:13–21.
- Glass I, Wollberg Z (1983b) Responses of cells in the auditory cortex of awake squirrel monkeys to normal and reversed species-specific vocalizations. *Hear Res* 9:27–33.
- Gourévitch B, Eggermont JJ (2007) Spatial representation of neural responses to natural and altered conspecific vocalizations in cat auditory cortex. *J Neurophysiol* 97:144–158.
- Grace JA, Amin N, Singh NC, Theunissen FE (2003) Selectivity for conspecific song in the zebra finch auditory forebrain. *J Neurophysiol* 89:472–487.
- Harper LV (1976) Behavior. In: *The biology of the guinea-pig* (Wagner JE, Manning PJ, eds), pp 31–51. New York: Academic.
- Heil P (1997a) Auditory cortical onset responses revisited. I. First-spike timing. *J Neurophysiol* 77:2616–2641.
- Heil P (1997b) Auditory cortical onset responses revisited. II. Response strength. *J Neurophysiol* 77:2642–2660.
- Heil P, Irvine DR (1996) On determinants of first-spike latency in auditory cortex. *Neuroreport* 7:3073–3076.
- Heil P, Irvine DR (1998) The posterior field P of cat auditory cortex: coding of envelope transients. *Cereb Cortex* 8:125–141.
- Hennevin E, Huetz C, Edeline JM (2007) Neural representations during sleep: from sensory processing to memory traces. *Neurobiol Learn Mem* 87:416–440.
- Huetz C, Del Negro C, Lebas N, Tarroux P, Edeline JM (2006) Contribution of spike timing to the information transmitted by HVC neurons. *Eur J Neurosci* 24:1091–1108.
- Huetz C, Philibert B, Guédin M, Avillac, M, Edeline J-M (2007) Spike timing carries information about conspecific vocalizations in the thalamo-cortical auditory system of guinea-pigs. *Soc Neurosci Abstr* 33:505.4.
- Janata P, Margoliash D (1999) Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *J Neurosci* 19:5108–5118.
- Kara P, Pezaris JS, Yurgenson S, Reid RC (2002) The spatial receptive field of thalamic inputs to single cortical simple cells revealed by the interaction of visual and electrical stimulation. *Proc Natl Acad Sci U S A* 99:16261–16266.
- Kumbhani RD, Nolt MJ, Palmer LA (2007) Precision, reliability, and information-theoretic analysis of visual thalamocortical neurons. *J Neurophysiol* 98:2647–2663.
- 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.
- Lewicki MS, Arthur BJ (1996) Hierarchical organization of auditory temporal context sensitivity. *J Neurosci* 16:6987–6998.
- Liu RC, Schreiner CE (2007) Auditory cortical detection and discrimination correlates with communicative significance. *PLoS Biol* 5:e173.
- Machens CK, Wehr MS, Zador AM (2004) Linearity of cortical receptive fields measured with natural sounds. *J Neurosci* 24:1089–1100.
- Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503–1506.
- Manunta Y, Edeline JM (1999) Effects of norepinephrine on frequency tuning of auditory cortex neurons during wakefulness and slow-wave sleep. *Eur J Neurosci* 11:2134–2150.
- Manunta Y, Edeline JM (2004) Noradrenergic induction of selective plasticity in the frequency tuning of auditory cortex neurons. *J Neurophysiol* 92:1445–1463.
- Margoliash D (1986) Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *J Neurosci* 6:1643–1661.
- Margoliash D (1997) Functional organization of forebrain pathways for song production and perception. *J Neurobiol* 33:671–693.
- Massaux A, Dutrieux G, Cotillon-Williams N, Manunta Y, Edeline JM (2004) Auditory thalamus bursts in anesthetized and non-anesthetized states: contribution to functional properties. *J Neurophysiol* 91:2117–2134.
- Mechler F, Victor JD, Purpura KP, Shapley R (1998) Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli. *J Neurosci* 18:6583–6598.
- Mickey BJ, Middlebrooks JC (2003) Representation of auditory space by cortical neurons in awake cats. *J Neurosci* 23:8649–8663.
- Middlebrooks JC, Clock AE, Xu L, Green DM (1994) A panoramic code for sound location by cortical neurons. *Science* 264:842–844.
- Middlebrooks JC, Xu L, Eddins AC, Green DM (1998) Codes for sound-source location in nontopographic auditory cortex. *J Neurophysiol* 80:863–881.
- Nakahara H, Zhang LI, Merzenich MM (2004) Specialization of primary auditory cortex processing by sound exposure in the “critical period.” *Proc Natl Acad Sci U S A* 101:7170–7174.
- Narayan R, Graña G, Sen K (2006) Distinct time scales in cortical discrimination of natural sounds in songbirds. *J Neurophysiol* 96:252–258.
- Nealen PM, Schmidt MF (2002) Comparative approaches to avian song system function: insights into auditory and motor processing. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 188:929–941.
- Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14:474–480.
- Nemenman I, Lewen GD, Bialek W, de Ruyter van Steveninck RR (2008) Neural coding of natural stimuli: information at sub-millisecond resolution. *PLoS Comput Biol* 4:e1000025.
- Pelleg-Toiba R, Wollberg Z (1991) Discrimination of communication calls in the squirrel monkey: “call detectors” or “cell ensembles”? *J Basic Clin Physiol Pharmacol* 2:257–272.
- Philibert B, Laudanski J, Edeline JM (2005) Auditory thalamus responses to species-specific guinea-pig vocalizations: a comparison between rat and guinea-pig. *Hear Res* 209:97–103.
- Populin LC (2005) Anesthetics change the excitation/inhibition balance that governs sensory processing in the cat superior colliculus. *J Neurosci* 25:5903–5914.
- Prather JF, Mooney R (2004) Neural correlates of learned song in the avian



- forebrain: simultaneous representation of self and others. *Curr Opin Neurobiol* 14:496–502.
- Reinagel P, Reid RC (2000) Temporal coding of visual information in the thalamus. *J Neurosci* 20:5392–5400.
- Reinagel P, Reid RC (2002) Precise firing events are conserved across neurons. *J Neurosci* 22:6837–6841.
- Rose HJ, Metherate R (2005) Auditory thalamocortical transmission is reliable and temporally precise. *J Neurophysiol* 94:2019–2030.
- Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. *Nat Neurosci* 1:513–518.
- Schnupp JW, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci* 26:4785–4795.
- Schreiber S, Whitmer D, Fellous JM, Tiesinga P, Sejnowski TJ (2003) A new correlation-based measure of spike timing reliability. *Neurocomputing* 52:925–931.
- Shannon RV, Zeng FG, Kamath V, Wyganski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304.
- Tiesinga P, Fellous JM, Sejnowski TJ (2008) Regulation of spike timing in visual cortical circuits. *Nat Rev Neurosci* 9:97–107.
- Tortorolo P, Falconi A, Morales-Cobas G, Velluti RA (2002) Inferior colliculus unitary activity in wakefulness, sleep and under barbiturates. *Brain Res* 935:9–15.
- Usrey WM, Reid RC (1999) Synchronous activity in the visual system. *Annu Rev Physiol* 61:435–456.
- Victor J, Purpura K (1997) Metric-space analysis of spike trains: theory, algorithms, and application. *Network* 8:127–164.
- Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J Neurophysiol* 76:1310–1326.
- Wallace MN, Palmer AR (2008) Laminar differences in the response properties of cells in the primary auditory cortex. *Exp Brain Res* 184:179–191.
- Wallace MN, Rutkowski RG, Palmer AR (2000) Identification and localisation of auditory areas in guinea pig cortex. *Exp Brain Res* 132:445–456.
- Wang L, Narayan R, Graña G, Shamir M, Sen K (2007) Cortical discrimination of complex natural stimuli: can single neurons match behavior? *J Neurosci* 27:582–589.
- Wang X (2000) On cortical coding of vocal communication sounds in primates. *Proc Natl Acad Sci U S A* 97:11843–11849.
- Wang X, Kadia SC (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 86:2616–2620.
- Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74:2685–2706.
- Wehr M, Zador AM (2003) Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426:442–446.
- Whitfield IC (1968) The medial geniculate body of the cat. *J Sound Vib* 8:108–117.
- Whitfield IC, Purser D (1972) Microelectrode study of the medial geniculate body in unanaesthetized free-moving cats. *Brain Behav Evol* 6:311–322.
- Wollberg Z, Newman JD (1972) Auditory cortex of squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science* 175:212–214.
- Wright BD, Sen K, Bialek W, Doupe AJ (2002) Spike timing and the coding of naturalistic sounds in a central auditory area of songbirds. In: *Advances in neural information processing systems*, Vol 14 (Dietterich TG, Becker S, Ghahramani Z, eds), pp 309–316. Cambridge, MA: MIT.
- Zhang LI, Tan AY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature* 424:201–205.