Spatial Representation of Neural Responses to Natural and Altered Conspecific Vocalizations in Cat Auditory Cortex

Boris Gourévitch and Jos J. Eggermont
Department of Physiology and Biophysics and Department of Psychology, University of Calgary, Calgary, Alberta, Canada

Submitted 3 August 2006; accepted in final form 29 September 2006

Gourévitch B, Eggermont JJ. Spatial representation of neural responses to natural and altered conspecific vocalizations in cat auditory cortex. J Neurophysiol 97: 144–158, 2007. First published October 4, 2006; doi:10.1152/jn.00807.2006. This study shows the neural representation of cat vocalizations, natural and altered with respect to carrier and envelope, as well as time-reversed, in four different areas of the auditory cortex. Multiunit activity recorded in primary auditory cortex (AI) of anesthetized cats mainly occurred at onsets (<200-ms latency) and at subsequent major peaks of the vocalization envelope and was significantly inhibited during the stationary course of the stimuli. The first 200 ms of processing appears crucial for discrimination of a vocalization in AI. The dorsal and ventral parts of AI appear to have different roles in coding vocalizations. The dorsal part potentially discriminated carrier-altered meows, whereas the ventral part showed differences primarily in its response to natural and time-reversed meows. In the posterior auditory field, the different temporal response types of neurons, as determined by their poststimulus time histograms, showed discrimination for carrier alterations in the meow. Sustained firing neurons in the posterior ectosylvian gyrus (EP) could discriminate, among others, by neural synchrony, temporal envelope alterations of the meow, and time reversion thereof. These findings suggest an important role of EP in the detection of information conveyed by the alterations of vocalizations. Discrimination of the neural responses to different alterations of vocalizations could be based on either firing rate, type of temporal response, or neural synchrony, suggesting that all these are likely simultaneously used in processing of natural and altered conspecific vocalizations.

INTRODUCTION

Conspecific communication sounds play an essential role in an animal’s behavior, influencing its social interactions, reproductive success, and survival. The wide vocal repertoire of mammals expresses a range of emotions conveyed by different calls or by alterations of a common call in the time–frequency domain. As part of the cat’s vocal repertoire, the “meow” is basically composed of harmonically structured amplitude-modulated sounds that are highly stereotyped from cat to cat, suggesting that the meow allows discrimination of the emitter. The meow is associated with demands, anger, and complaints (Moelk 1944), and its emotional salience might be conveyed by the temporal aspects, whereas the harmonic content might be crucial for identification of the vocalizing animal.

The neural representation of vocalizations begins at the level of the auditory nerve with temporal and rate-based codes for basic spectrotemporal features of the stimulus (Delgutte and Kiang 1984; Sachs and Young 1979; Young and Sachs 1979). Some neural subpopulations of the cochlear nucleus preserve or enhance the codes stemming from auditory nerve fibers (Blackburn and Sachs 1990). The cortical processing of vocalizations is less clear. At one time, it was thought that specific neurons named “call detectors” might be specifically responsive to particular vocalizations. After several reports of such neurons (Funkenstein and Winter 1973; Manley and Muller-Preuss 1978; Newman and Wollberg 1973; Winter and Funkenstein 1973; Wollberg and Newman 1972), it became increasingly apparent that these neurons could respond to numerous call types or only to complex spectral components thereof. Currently, it is generally accepted that primary auditory cortex (AI) encodes acoustic features of vocalizations through frequency decomposition, whereas encoding of temporal aspects requires higher-order processing. Thus identification of more complex features or emotions associated with the stimulus could be realized in lateral belt areas, or their equivalent in the cat, or in prefrontal cortical areas (Rauschecker and Tian 2000; Romanski et al. 2005; Schnupp et al. 2006).

Higher-order processing could be realized either by specialized neurons or/and by discharge patterns of neuronal assemblies. This last hypothesis is very difficult to prove, at least with the current recording techniques involving generally a small number of units; thus the focus is largely on the properties of single neurons or neuron clusters responding to complex sounds. In general, because vocalizations are a mixture of amplitude- and frequency-modulated harmonic components, numerous neurons respond in a combination-sensitive manner, i.e., the addition of responses to parts of the stimulus is nonlinear (Rauschecker et al. 1995, 1998) and either larger than the response to the whole stimulus or smaller (Gehr et al. 2000; Rauschecker et al. 1995).

The quest for call detectors has never really ended. For that purpose, numerous studies have compared natural and time-reversed vocalizations (Gehr et al. 2000; Glass and Wollberg 1983; Pelleg-Toiba and Wollberg 1991; Wang et al. 1995). Although the different temporal acoustical features of such time-reversed stimuli may induce biased results, they are of interest to understand the global identification of conspecific-stimulus relevance (Wang and Kadia 2001). These studies generally concluded that the overall firing rate (FR) was not significantly modified by time reversal of the vocalization. However, a more detailed mapping of neurons showing differences in the response to forward and time-reversed stimuli is currently not available for the cat auditory cortex.

Our previous study showed the inefficiency, albeit not the impossibility, of a firing rate–based discrimination between alterations of vocalizations in AI (Gehr et al. 2000). The current study investigates the discrimination abilities of several cortical areas of the cat and presents cortical maps of neurons...
with abilities to discriminate some temporal, carrier, or time-reversing alterations of the natural meow. The main results point to the importance of the posterior ectosylvian gyrus (EP) areas in the representation of the information conveyed by the alterations of vocalizations and to the different roles of dorsal and ventral parts of AI. We also show that several parameters including FR, type of temporal response pattern, and neural synchrony can be used to discriminate the alterations of cat meows, suggesting a simultaneous use of several codes in vocalization processing rather than one unique code.

METHODS

The care and the use of animals reported in this study was approved (#B1 2001–021) and reviewed on a yearly basis by the Life and Environmental Sciences Animal Care Committee of the University of Calgary. All animals were maintained and handled according to the guidelines set by the Canadian Council of Animal Care.

Animal preparation

All animals were deeply anesthetized with the administration of 25 mg/kg of ketamine hydrochloride and 20 mg/kg of sodium pentobarbital, injected intramuscularly. A mixture of 0.2 ml of acepromazine (0.25 mg/ml) and 0.8 ml of atroine methyl nitrate (25 mg/ml) was administered subcutaneously (sc) at approximately 0.25 ml/kg body weight. Lidocaine (20 mg/ml) was injected sc before incision. The tissue overlying the right temporal lobe was removed and the dura was resected to expose the area bounded by anterior and posterior ectosylvian sulci. The cat was then secured with one screw cemented on the head without any other restraint. The wound margins were infused every 2 h with lidocain and additional acepromazine/atropine mixture was administered every 2 h. The temperature of the cat was maintained at 37°C. The ketamine dose to maintain a state of areflexive anesthesia in this study was on average 9.2 mg · kg⁻¹ · h⁻¹ (range 6–13 mg · kg⁻¹ · h⁻¹).

Acoustic stimulus presentation

Stimuli were generated in MATLAB and transferred to the DSP boards of a TDT-2 (Tucker Davis Technologies) sound-delivery system. Acoustic stimuli were presented in an anechoic room from a speaker system (Fostex RM765 in combination with a Realistic Super-Tweeter that produced a flat spectrum (≤5 dB) ≤40 kHz measured at the cat’s head) placed about 30° from the midline into the contralateral field, about 50 cm from the cat’s left ear. Calibration and monitoring of the sound field was accomplished with a condenser microphone (Brüel & Kjér 4134) placed above the animal’s head, facing the speaker, and a measuring amplifier (Brüel & Kjér 2636). Before acute recordings peripheral hearing sensitivity was determined using auditory brain stem response (ABR) thresholds (details in Norena et al. 2003).

Frequency-tuning curves were measured by randomly presenting 27 gamma-tone pips with frequencies covering 5 octaves (e.g., 1.25–40 kHz) in equal logarithmic steps and presented at eight different stimulus levels in 10-dB steps (e.g., 5- to 75-dB SPL) at a rate of four per second such that each intensity-frequency combination was repeated five times. The envelope of the gamma tones is given by

\[ y(t) = (t/4)^2 \exp(-t/4) \]

with t in milliseconds. The duration of the gamma tones at half-peak amplitude was 15 ms and the envelope was truncated at 50 ms, where the amplitude is down by 64 dB compared with its peak value. Best frequency (BF) of the individual recordings was determined at 65-dB SPL.

Then, a typical kitten’s meow (Brown et al. 1978; Romand and Ehret 1984) was used as a natural stimulus and altered in its temporal and spectral properties. The intensity level of presentation was 65 dB, i.e., at the same level as used to determine BFs. The stimulus set was identical to that used in our previous study (Gehr et al. 2000). The duration of the natural meow was 0.87 s, the average fundamental frequency was 570 Hz, the lowest frequency component was 0.5 kHz, and the highest frequency component was 5.2 kHz. The second and the third harmonics, between 1.5 and 2.5 kHz, had the highest intensity. Distinct frequency modulations occurred simultaneously in all formants between 100 and 200 ms after onset. The meow was also presented in a factor 1.5 time-expanded form and in a factor 0.75 time-expanded (i.e., compressed) form while keeping the frequency spectrum the same. The spectral contents of these three stimulus types were also increased by a factor 2 and lowered by a factor 2. Therefore a set of nine stimuli finally resulted from these morphing procedures (Fig. 1). The temporally and spectrally transformed meows were computed using Lemur software (CERL Sound Group, University of Illinois).

The low-frequency meows have very little energy above 2.5 kHz and the high-frequency meows have most of their energy above 2.5 kHz. The three time-expanded stimuli were presented before the three natural-length stimuli and before the three time-compressed stimuli. For each type, the low-frequency stimulus was presented first, followed by the natural-frequency stimulus and the high-frequency stimulus. This stimulus set was presented to the animal in both normal and time-reversed form. The individual meow stimuli were repeated 25 times. Stimuli were presented once per 3 s.

Recording

Two arrays of eight electrodes (FHC) each with impedances between 1 and 2 MΩ were used. The electrodes were arranged in a 4 × 2 configuration with interelectrode distance within rows and columns equal to 0.5 mm. Each electrode array was oriented such that all electrodes were touching the cortical surface and then were manually and independently advanced using a Narishige M101 hydraulic microdrive (one drive for each array). The depth of recording was set to 600 and 1,200 μm and thus the electrodes were likely in deep layer III or layer IV. The signals were amplified 10,000 × using the FHC HiZx8 set of amplifiers with filter cutoff frequencies set at 300 Hz and 5 kHz. The signals were processed by a TDT-Pentusa multichannel data-acquisition system (filter bandwidth 300 Hz to 10 kHz). Spike sorting was done off-line using a semiautomated procedure based on principal component analysis and K-means clustering implemented in MATLAB. The spike times and waveforms were stored. The multiple single-unit data presented in this paper represent only well-separated single units that, because of their regular spike wave form, likely are dominantly from pyramidal cells, thereby eliminating potential contributions from thalamocortical afferents or fast spikes from interneurons. For statistical purposes, the separated single-unit spike trains were added again to form a multieit spike train.

Data analysis

FIRING RATE. For each site, the multiunit firing rate (FR) associated with a specific stimulus was defined as the number of spikes recorded by the electrode divided by the duration of the stimulus. For each site, the comparison of the FR between two stimuli chosen from the 18 stimuli previously presented or during 15 min of silence was quantified by the change in decibels (dB), defined as \(20 \times \log_{10}(\text{FR2}/\text{FR1})\), FR1 and FR2 being the FR induced by the two stimuli, respectively (+6 dB thus indicates that FR2 is twice FR1). The one-sided Wilcoxon test (Wilcoxon 1945) was used to compare the FRs on all sites produced by the presentation of two stimuli. In all other cases, the
one-sided Mann–Whitney test (Mann and Whitney 1947) was used to compare two unpaired sets of values.

SPATIAL MAPPING. Four cortical fields were studied: AI, anterior auditory field (AAF), posterior auditory field (PAF), and both dorsal and intermediate parts of the posterior ectosylvian gyrus (EP) (Read et al. 2002). The distinction between cortical fields AI and AAF was based on the frequency gradient along the multielectrode recording array and the width of the frequency-tuning curve bandwidth at 20 dB above threshold. The demarcation of the AI and the posterior fields was based on response latency, clearness of frequency tuning, and nonmonotonicity. Thus long latencies, narrow frequency tuning, and strongly nonmonotonic rate intensity functions were used to assign neurons to PAF, whereas long latencies and fuzzy or absent frequency tuning assigned recordings to EP. We initially analyzed the dorsal and intermediate parts of EP separately, but because we did not obtain differences for any of the parameters studied we combined the data.

To allow comparison between the localizations of the sites on different cats, the distance between the tips of the AES (anterior ectosylvian sulcus) and PES (posterior ectosylvian sulcus) was normalized to 100% on each cat, then a coordinate for each site was computed as follows: the abscissa is the percentage of the AES–PES distance, 0 being the tip of the PES and 100 the tip of the AES; the ordinate is the distance (in millimeters) in the ventrodorsal direction from the tip of the PES, perpendicularly to the PES–AES axis (Norena et al. 2006). The coordinates of the sites were finally superimposed on a representative image of the cat auditory cortex (Fig. 2). The data points mapped are from those sites that had an average FR $>3$ spikes/s in response to all stimuli.

PSTH CLASSIFICATION. For each recording site, the multiunit post-stimulus time histogram (PSTH) was constructed using a bin width of 5 ms. Even if a continuous distribution of neural responses is often observed (Campbell et al. 2006), we may assume that different neural responses are underlying different processing. We thus classified the PSTH types observed as follows: after being smoothed with a three-bin running average, the bins with counts $>50\%$ of the maximal bin were retained and all retained bins separated by $>20$ ms were merged. Each merged set is then considered as a “peak” in the PSTH. We defined three types of PSTH based on the number of peaks: 1) “Onset”: all the peaks occur before 160 ms after stimulus onset. 2) “FewPeaks”: in addition to onset peaks, three or fewer peaks occur after 160 ms. 3) “Sustained”: more than three peaks are observed after 160 ms.
160 ms (Fig. 3). In all studies, except if noted, the PSTH type refers to the response to a natural meow.

NEURAL SYNCHRONY. Cross-correlograms were calculated using custom-written programs in MATLAB (Eggermont and Smith 1995). The bin size was 2 ms and the resulting cross-correlogram was smoothed with a three-bin running average. We used the peak cross-correlation coefficient between neural responses to quantify neural synchrony. Stationary estimates of the recordings were based on firing rate (mean and variance) in 100-s-long segments of the 15-min recordings for silence and over 25 trials for the meow stimuli. To correct for the overall firing rate, burst firing, and periodicities in the firing of the neurons, the cross-covariance was deconvolved with the square root of the product of the autocovariance functions. This deconvolution was done in the frequency domain, where it becomes a simple division; Fourier transformation back to the time domain resulted in the corrected cross-correlation coefficient function. The temporal evolution of the synchrony during stimuli was studied by the mean of the synchrony computation on 29 sliding time windows of 100-ms width for the natural meows, 70 ms for the compressed meow, and 150 ms for the extended one. The time windows are regularly spaced to cover the whole stimulus length for each stimulus. Consequently, the fifth window applied on the natural meow corresponds to the same relative portion of stimulus as the fifth window applied on the compressed meow.

HIERARCHICAL CLUSTERING. The method of hierarchical clustering (Everitt 1978) applies when some clusters are nested within other clusters and the technique operates on a matrix of individual similarities or distances. For hierarchical clustering using firing rate and neural synchrony, we used the standardized Euclidean distance, where each coordinate in the sum of squares is inverse weighted by the sample variance of that coordinate. For clustering using PSTH types, we used the Hamming distance, i.e., the percentage of coordinates that differ. The implementation in MATLAB used the Ward single-linkage method based on a minimum-variance algorithm to form clusters, which are visualized in a dendrogram. An inconsistency level value of 1.0 was used.

RESULTS

We obtained data from 400 recording sites in 16 adult cats. Of the 400 sites recorded from 243, having an average FR >3 spikes/s in response to all stimuli, were kept for the study. Of the 243 sites analyzed, 182 sites were localized in AI (74.9%), 20 in AAF (8.2%), 20 in PAF (8.2%), and 21 in EP (8.6%).

Firing rate

NATURAL MEOW VERSUS SILENCE. The distribution of spontaneous FR for the different cortical areas is shown in Fig. 4A. The natural meow stimulus induces a significant global increase (Wilcoxon tests, \( P < 10^{-5} \)) of the FR compared with silence for Onset and FewPeaks neurons (Fig. 4B). However, the FR for the natural meow rarely exceeds twice the spontaneous FR. The results are highly dependent on the cortical area and the type of response: all neurons PSTH types in AI, but only Onset neurons in AAF and PAF, showed a clear increase in FR. Furthermore, FewPeaks neurons in the EP strongly responded to the natural meow, whereas Sustained neurons were inhibited (Fig. 4B). The AI and EP areas also showed the highest variability of increases in FR, indicating a strong involvement of these two areas in the response to vocaliza-
tions. Finally, all auditory areas included sites showing large increases or decreases (>6 dB) in FR to the meow relative to spontaneous activity (Fig. 4C).

CARRIER EFFECT. Changes in FR for natural-envelope meows with low- or high-frequency carriers compared with the natural meow were fairly modest on average. In AI, the modification of the carrier pitch has effects that are almost independent of the PSTH-type of the neuron (Fig. 5, A and C). This may be a consequence of the unchanged temporal envelope. As expected, we noticed a slight decrease in the FR for the low-frequency meow in the dorsoanterior part of the AI area (Fig. 5B), where neurons have high BFs (Fig. 2A). However, these same neurons also showed a decrease in FR for a high-frequency meow, suggesting that a more specialized processing may occur at this place. The high-frequency meow also induced an increase of FR just above the tip of the PES (Fig. 5D). These neurons were mostly assigned to PAF or EP (Fig. 5, A and C). The FewPeaks neurons of these two areas show opposite behaviors for low- and high-frequency meows, whereas the Sustained neurons of these areas show the same responses to both stimuli.

TEMPORAL ENVELOPE EFFECT. Next, we consider the changes in FR for the compressed or expanded meow with natural carrier compared with the natural meow. The FR hardly changed for compression of a meow and the neurons that did show FR differences were distributed widely over the cortex (Fig. 6B). In contrast, the expanded meow induced a general decrease in FR (Wilcoxon test, P < 0.05) for all types of neurons except Onset in AAF and PAF (Fig. 6C), but the most affected ones were dominantly localized in the dorsoanterior part of the AI area (Fig. 6D). A higher proportion of neurons of EP areas shows a modest (mostly <3 dB) increase in FR for both compressed and expanded meows compared with the average change in AI.

TIME-REVERSAL EFFECT. The differences in FR induced by time reversal of the natural meow for AI Onset or FewPeaks neurons were larger (Mann–Whitney test, P < 7 × 10^{-3}, P = 0.07, respectively) than for the Sustained neurons (Fig. 7A). This is likely a consequence of the longer ramp of the temporal envelope of the time-reversed meow. In contrast, the Sustained neurons in the EP areas seemed to show slightly larger, though not significant, differences than the FewPeaks neurons (Fig. 7A). A striking finding is the localization of the neurons that showed an increase in FR of >3 dB: about 60% of them were in the posterior part of the auditory cortex, i.e., in AI or EP, and only one in PAF. About 35% of them were in the ventral part of the AI (Fig. 7B).

Temporal response properties

PSTH TYPES. For the natural meow and among the 243 sites considered, 94 sites were classified as Onset type (38.7%), 89 as FewPeaks type (36.6%), and 60 were of the Sustained type (24.7%). For responses to a natural meow, the distribution of each neuron’s PSTH type along the PES–AES horizontal axis is presented in Fig. 8. The Sustained neurons are mainly found in the posterior part of the auditory cortex, whereas most Onset neurons are found in the caudal half (0–50%) between the ectosylvian sulci. The FewPeaks neurons are spread out more evenly over the PES–AES axis. AI has mainly Onset and FewPeaks neurons, whereas fewer Onset and more Sustained neurons are found in AAF and PAF. Interestingly, only FewPeaks and mainly Sustained neurons are found in EP (Fig. 8C). As expected, the peak latency of the PSTH falls in the first 120 ms and is generally equal to the latency of the first onset peak. It is shortest in AI and AAF, then in PAF, and longest in the EP areas (Fig. 8D).

COMPONENTS OF PSTHS. For the natural meow, peaks in the PSTH mainly occurred before 250 ms (Fig. 9A) irrespective of
the BF of the site. The peaks are slightly more spread out over time for sites with high BF. A large proportion of peaks is also visible around 750 ms for sites with BF < 2.5 kHz. Globally, vocalizations induced significantly stronger activity than silence until 250–300 ms after stimulus onset (Wilcoxon test, \( P < 0.05 \), Fig. 9B) and around 750 ms for low-BF sites. In contrast, the neurons are on average inhibited during the remainder of the stimulus (Wilcoxon test, \( P < 0.05 \), Fig. 9B), i.e., fire at a lower rate than during silence.

For the natural meow and the altered stimuli described so far, there was a correspondence between the BF of a site, the number and latency of the peaks in the PSTH of this site, and the frequency–time representation of the stimulus (Fig. 10). For all forward stimuli, the PSTHs generally showed an onset peak at about 30 ms followed by a rebound at about 100 ms. For neurons with BFs outside the frequency content of the stimulus, i.e., for frequencies > 5 kHz, onset and rebound responses were also present but more spread out over time. For all forward stimuli, a peak was also present around 750 ms for the natural, low-frequency, and high-frequency meows (corresponding to 550 ms and 1.1 s for compressed and expanded meow, respectively) and was limited to the BF range corresponding to the frequency content of the stimulus at this time. In other words, the neurons having a BF at a frequency not found in the stimulus at this time are scarcely or not responding. Except at these specific times, which may be crucial for vocalization identification, there were few peaks during the stimulus presentation, consistent with the time range of inhibition described in the previous section. For all forward stimuli, the neurons without a clear BF generally show a behavior similar to that for high-BF neurons, albeit fuzzier.

The time-reversed meow begins with a slow ramp in its temporal envelope. Consequently, the responses show a long-lasting onset and rebound. A peak is also visible at 850 ms for the neurons having a BF in the frequency range of the reverse meow at this time, corresponding to a big gap in the energy of the stimulus for frequencies < 2.5 kHz.

**Figure 6.** Changes in the FR between a natural meow stimulus and time-compressed (A, B) and time-expanded (C, D) meows. A and C: values for all sites and each auditory area. PSTH types are specified. B and D: mapping on the cortex of the extreme values of the changes in FR.

**Figure 7.** Change in the FR between a natural meow and a time-reversed meow. A: values for all sites and each auditory area. PSTH types are specified. B: mapping on the cortex of the extreme values of the increase in FR for the time-reversed meow.
(two first bins, compressed meow). For PAF and EP areas, an inhibition period immediately followed this response (bins 3 and 4, compressed meow). In AI, during the first half of the stimulus, the compressed meow provoked a stronger response, whereas the expanded meow generally induced a weaker response during almost the entire stimulus, especially after the third bin (150 ms of the natural meow stimulus). In contrast to the findings in AI, the behavior of both PAF and EP in the first half of the stimulus was quite similar for compressed and expanded meows, except in the two first bins in PAF. Compression and expansion had a less clear effect at the end of the stimulus, probably because of lower stimulus level. The AAF and AI showed similar differences in the patterns of time activations for both compressed and expanded meows compared with natural meow.

Time-reversed meows induced mostly inhibition in AAF. Contrary to findings in AI, it evoked similar behavior in PAF and EP areas as the expansion of the meow. Interestingly, all alterations of the meow induced a larger excitation of neurons in EP between about 200 and 500 ms compared with the response to a natural meow. However, the periods of strongest excitations had different time distributions in this interval. For the reversed and expanded stimuli, the percentages were close to the expected level of 50% at the beginning of the stimulus in AI. This means the FR is similar for the meow, the expanded meow, and the time-reversed meow at the very beginning of the response, even with the slower rising ramp of the two latter ones.

From these results, it is apparent that 1) there is a predominant onset response irrespective of whether the stimulus is time reversed; 2) EP and PAF neurons show temporal variations in the discrimination between a natural and a temporally altered vocalization, whereas AAF rather simply shows similar or opposite behaviors to AI; 3) in all areas, the presence of numerous large differences shows that the temporal alterations of the natural meow induce responses very different from those obtained by performing the same alterations on the response to a meow. If this was the case, given that the same portions of stimulus are compared between a natural meow and a compressed or expanded meow, then the differences shown in Fig. 11, A and B would need to be zero.

**Neural synchrony**

**NATURAL MEOW AND TIME-REVERSED MEOW.** During silence, the strongest synchrony is found between neighboring electrodes and decreases with distance between recording sites.
The highest values occur between 10 and 110% of the PES–AES distance, i.e., within AI. In response to a meow, the synchrony is globally increased all along the horizontal posterior–anterior axis and between sites at any distance (Fig. 12B). The evolution of the average synchrony with time since stimulus onset in each auditory area is shown for a comparison between a natural meow and silence (Fig. 13A) and for the time-reversed natural meow compared with the natural meow (Fig. 13B).

On average, the natural meow causes a higher synchrony between sites within or between cortical areas compared with silence. Within AI, except during the rebound response around 150 ms, the synchrony remains at the same level during the entire stimulus duration (1 s shown). Interestingly, the highest increase in synchrony between sites in AI and EP is found during the onset, whereas the highest one within EP areas sites occurs between 300 and 500 ms. This last result might indicate a crucial point in time for high-order processing of vocaliza-
tions in EP areas. In other areas, the relations between neurons are not clearly modified during time since stimulus onset. Similarly, the strongest differences in average synchrony between a forward and a time-reversed meow are observed within EP area sites, whereas no apparent difference is visible in AI. The negative synchrony difference between 300 and 500 ms within EP may reflect a crucial period for vocalization processing (Fig. 14B) because this is a region where there is a large difference in synchrony between the meow and silence (Fig. 14A) and there is also a large excitation of neurons when the temporal envelope of the meow is altered or reversed (Fig. 11). Compared with the natural meow, the sites in EP consistently show more synchrony at the end of the time-reversed meow, i.e., during its highest energy periods. In contrast with EP, the synchrony differences between AI, AAF, and PAF are mostly positive. They also are mainly restrained to the time interval of 300–800 ms.

EFFECTS ON SYNCHRONY BY CHANGES IN CARRIER OR ENVELOPE. Envelope changes induced massive shifts in synchrony: the compressed meow provoked stronger correlations between all neurons during the entire stimulus (Fig. 14A). The expanded meow induced synchrony between 150 and 300 ms and also stronger synchrony otherwise (Fig. 14B). For this latter stimulus, the highest increase in synchrony during time involves sites with BF >10 kHz. In contrast, changes in the carrier frequency of the meows did not induce large changes on the synchrony between sites, irrespective of the BF of the site or the portion of stimulus (Fig. 14, C and D).

Similarities in responses to natural and altered meows

Obvious differences between the FR, the PSTH types, and the synchrony involved in, e.g., a response to a forward meow and a time-reverse meow can be elucidated by a hierarchical clustering (Fig. 15). Precisely, for clustering based on FR or PSTH type, all the time-reversed meows were farther from the natural meow than any alteration in envelope or carrier of the forward meow. This was also true for clustering based on synchrony except for the time-reversed compressed and the forward expanded meows. Interestingly, for the FR and the PSTH types, but not for synchrony, the alteration of the envelope induced fewer changes than the carrier modification (expanded, compressed, and natural are most often classified together before being lumped with low-frequency or high-frequency stimuli).

DISCUSSION

Limits of the study

Preliminary to the general discussion, we want to emphasize that the study deals with the response to vocalizations under ketamine anesthesia. Its effects on vocalization processing have been shown to be sometimes dramatic, at least in guinea pigs, with mostly suppressive effects in the temporal patterns of response to vocalizations and some strengthened onsets (Syka et al. 2005). We also emphasize a bias in the selection of responses to analyze, which required that they all had a FR >3 spikes/s. This may leave out potentially important neurons or recording sites. DeWeese et al. (2003) found that patch-clamped neurons in auditory cortex (under ketamine anesthesia) fired either 0 or 1 spike to tone bursts. These findings suggest a different selection bias for standard extracellular recordings and patch-clamp recordings and they also suggest that the number of units contributing to a sort output should be larger than assumed. In awake paralyzed cats, 40% of neurons in AI had spontaneous firing rates <1 spike/s (Aboeles and Goldstein 1970). If the number of contributing units is underestimated, even well-separated single-unit recordings must then be composed of activity from ≤10 not-well-synchronized units (otherwise the spikes would be superimposed and likely sorted as one unit). Thus leaving out low-firing multiunit clusters could have an impact on our conclusions. However, for our PSTH typing it was necessary to consider sites with higher firing rates.

FIG. 12. Neural synchrony maps in auditory cortex. Synchrony, defined here as the peak strength of the corrected cross-correlogram (METHODS), is plotted here as a function of the position of the recording electrode along the posteroanterior axis (abscissa) and the distance between the 2 electrodes involved in the calculation of synchrony (ordinate), for silence (A) and meow stimuli (B). For each electrode pair, positions along the posteroanterior axis are plotted. Color bar indicates strength of neural synchrony.

FIG. 13. Comparison of differences in synchrony within and between areas over time since stimulus onset between silence and stimulation with the natural meow (A) and between the time-reversed and natural meow (B). See METHODS for details.
Another limitation of the study is the small sample size of data in cortical areas other than AI. This was mainly a result of the more difficult access to AAF, PAF, and EP areas with our microelectrode arrays, the unclear anatomical borders between PAF, and EP combined with the weak responses of some electrodes in these areas. As a consequence, AAF, PAF, and EP areas often show a set of responses subject to high variance, which might limit conclusions for these nonprimary auditory cortical areas.

Assignment of recording sites to cortical areas

The natural variability of the location of cortical areas with respect to anatomical landmarks across cats induces some partial overlapping of the areas on the composite spatial mapping that we used. Nevertheless, we think the composite maps emphasize the trends in localization of the cortical processing of meows across cats. The assignment of recording sites to PAF and EP was difficult and the map of assigned recording sites (Fig. 2B) shows considerable overlap between units assigned to PAF and those to EP. Based on anatomical conditions alone the EP assignments would have been dorsal to PAF (Read et al. 2002). In individual animals this was the case. However, functional criteria such as long latencies, clear frequency tuning, and strongly nonmonotonic rate-intensity functions assigned neurons to PAF, whereas long latencies and fuzzy or absent frequency tuning indicated recordings from EP.

FIG. 14. Changes in synchrony with time after stimulus onset after envelope or carrier changes of the meow. Percentage of sites having a higher synchrony under compressed (A), expanded (B), low frequency (C), and high frequency meow (D) than under a natural meow compared to the expected percentage of 50% (for no effect). Each percentage is computed for a bin of 100-ms duration and 1-octave BF difference. For the compressed and expanded meows, the time intervals were chosen to correspond to the same relative portions of stimulus for the natural meow (see METHODS).

FIG. 15. Hierarchical clustering showing similarities in responses to the 18 stimuli based on (A) the type of response (percentage of sites showing a different type of PSTH between 2 stimuli), (B) the firing rate (Euclidean distance on FR normalized by the SD of each coordinate to reduce the effect of FR differences between sites), and (C) synchrony (same distance as for FR). R, reverse; N, natural; L, low frequency; H, high frequency; C, compressed; E, expanded; S, silence. Gray and hatched parts emphasize the distinction between time-reversed (gray) and forward (diagonal hatching) meows. Ward linkage rule is used for all classifications.
A problem is that there are, to our knowledge, no reports about the functional properties of EP neurons. Thus it is possible that our recordings in the posterior part of cortex may mostly have been from EP and, in that case, EP would be composed of neurons with response properties not unlike those reported for PAF as well as others not found in PAF. What is obvious is that the functional separation based on sharpness of frequency tuning and nonmonotonicity plays a distinguishing role in the response to the time-reversed meow sound.

Spectrotemporal response to a natural meow

Several studies have pointed to a strong relation between the neural response and the envelope of the vocalization band-pass filtered around the characteristic frequency of the neuron (Gehr et al. 2000; Wang 2000). We confirmed this by the systematic study of the correspondence between the time frequency representation of the vocalization, the BF of sites, and the peaks in the PSTH of these sites (Fig. 10). We noticed indeed that a peak around 750 ms was generated by neurons having a BF < 2.5 kHz, i.e., in the range of frequencies composing the stimulus at this time (Fig. 10A). However, most of the peaks in the PSTH occurred before 200 ms, when all distinct frequency modulations have already occurred (Fig. 1). This 200-ms period after stimulus onset showed the strongest neural activity, whatever the BF of the site, and was followed by a period of inhibition of the FR to levels significantly below the spontaneous FR during the rest of the stimulus (Fig. 9). This result is in agreement with the study of Sovijarvi (1975) showing that cat vocalizations excited cells in AI largely at the onset and offset of the stimulus and caused inhibition or no response at all during the other parts of the sound. The result also corroborates our earlier study (Gehr et al. 2000). This phenomenon was apparently present in other studies (Fig. 3 in Wang et al. 1995; Fig. 4 in Schnupp et al. 2006), although not discussed. The statistical significance of the inhibition of FR (Fig. 9) would indicate that more neurons show an ON-OFF (Onset/peak around 750 ms) behavior than a steady-state behavior. The first 200 or 250 ms after stimulus onset therefore appears crucial for the processing of the vocalization in AI given that afterward, the activity is decreasing. In EP areas, the synchrony is found to be stronger during the 300- to 500-ms period (Fig. 13).

Call detectors revisited

Call detectors are generally considered as hypothetical neurons able to encode a specific call or vocalization, often presenting a sustained or increased activity in response to the call. Their existence would support the hypothesis of a specialization of neurons for very complex sounds. The numerous candidates found in the earlier studies on primates were in fact shown to respond to more than one call or to various features of calls (Funkenstein and Winter 1973; Manley and Muller-Preuss 1978; Newman and Wollberg 1973; Winter and Funkenstein 1973; Wollberg and Newman 1972). It now appears that such neurons are rare or even absent in the early cortical processing stages (Wang et al. 1995). Consequently, the hypothesis has been controversial since then, even if some recent studies continue to question it. For example, Rauschecker and Tian (2000) found some neurons in the lateral belt of primates responding to total calls with the full spectrum but not as well to the low-pass–filtered version and not at all to the high-pass–filtered version.

Another way to identify call detectors is to study the activity induced by a time-reversed call or vocalization compared with the natural vocalization (Gehr et al. 2000; Glass and Wollberg 1983; Pelleg-Toiba and Wollberg 1991; Wang et al. 1995). These studies generally concluded that the global FR was not significantly modified by the reversing of the vocalization but there were some unequivocal temporal differences in the processing of the responses (Wang et al. 1995). In the study of Wang et al. on monkey calls, however, the high variability (about 8 Hz) and asymmetry of the temporal envelope of the spectral components of the call might also explain some parts of this result. These studies generally concluded that complex sounds were represented in the auditory cortex by the synchronized activity of functional cell ensembles because the time distribution of response peaks closely approximated in time with the envelope of a particular spectral component of the call, corresponding with the cell’s best frequency. We agree with this hypothesis, at least as far as AI is concerned. However, in our study, the strongest difference of FR between the forward and time-reversed vocalizations is observed in the EP areas (Fig. 7B) that constitute units showing the FewPeaks or Sustained PSTH response types and that are less sensitive to onsets, thereby emphasizing the difference of envelope between the two stimuli (Fig. 8C), although this preference for the forward vocalization is not a general behavior for all neurons with sustained responses (Fig. 7A). Some Onset type neurons in AI also show strong differences of FR between the forward and reversed meows (Fig. 7, A and B). Moreover, as several studies pointed out, in the guinea pig numerous response behaviors to vocalizations are observed in the AI area (Wallace et al. 2005) and in the inferior colliculus (Suta et al. 2003).

In our study, the location of the most active neurons in the forward-reverse discrimination, i.e., of the possible call detectors, appears to be clearly in the posterior part of the auditory cortex (EP) or the ventral part of AI, perhaps continuing in the dorsal part of secondary auditory cortex (AII) area not studied here (Fig. 7C). This is consistent with studies of conspecific vocalizations in primates showing that call specificity, manifest in coarse overall changes in firing rate, may be more common in higher-order cortical fields, i.e., lateral belt areas (Rauschecker and Tian 2000).

An original way to validate the consistency of experiments involving a forward stimulus and its time-reversed version is to present to an animal a call from another animal and its time-reversed version. Responses from primary auditory cortex in marmosets and cats responding to a twitter call from a marmoset showed that the marmoset was better able than the cat to discriminate the natural call than the time-reversed version (Wang and Kadia 2001). This would suggest that the differences observed between responses to forward and time-reversed stimuli stem from difference in relevance and not only from the differences in acoustical structures of the two stimuli.

Effect of the carrier alteration of the meow

On average, AI and EP showed the largest FR changes in the comparison between natural meows and meows with altered carrier frequency. The Sustained type neurons in EP mostly
show an increase whatever the alteration, whereas the response of FewPeaks neurons was more discriminating between the low-frequency and high-frequency alterations. Amazingly, each neuron type in PAF showed the opposite behavior in response to low-frequency and high-frequency alterations (Fig. 5, A and C). This might emphasize the different role of EP and PAF in identification of vocalizations or their emitters. Nearly half of the sites showing the strongest decrease in FR after a carrier alteration are localized in the dorsal part of the AI area (Fig. 5, A and C). The slight decrease of FR induced by the low-frequency meow compared with the natural meow may probably be explained by the restricted bandwidth of the low-frequency meow, involving fewer neurons in the response. For this stimulus, we also observed an absence of strong rebounds after the onset in the high-BF sites (Fig. 10C). The opposite phenomenon occurs in response to the high-frequency meow. Finally, carrier alterations appear to have little effect on synchrony (Fig. 14, C and D).

Effect of the temporal envelope alteration of the meow

Several studies showed some specificity in AI for small changes in temporal envelope: the rise time of a stimulus is known to have an effect on latency and strength of the neuronal response (Fishbach et al. 2001; Heil 1997a,b; Phillips 1998). Temporal asymmetry in ramped and damped sinusoids with a short period (25 ms) was clearly reflected in average discharge rate differences but not necessarily by temporal discharge patterns of auditory cortical neurons in primates (Lu et al. 2001). For bird chirps, every small temporal perturbation (denoising, artificial stimulus with same time frequency features) was shown to have a substantial influence on the responses in cat AI (Bar-Yosef et al. 2002). Here, we consistently observed some local differences in the FR induced by envelope alteration of the meow (Fig. 11). More precisely, the alteration of the envelope of the meow rarely produced PSTHs to those that resulted from the same alteration of the PSTHs obtained for stimulation with the natural meow.

This suggests that information is associated with temporal alterations of the forward vocalizations. The code for such alterations would be strictly temporal or strongly localized because expanded or compressed stimuli induce global FR and PSTH types that are more similar to those recorded under a natural meow than carrier-altered stimuli (Fig. 15). There also may be a synchronization aspect of such a code because synchronization was much more sensitive to temporal alterations compared with carrier alterations (Fig. 14).

The mapping of sites showing a decrease in FR for the expanded meow compared with the natural meow strongly involves the AI in the temporal processing of the expanded meow (Fig. 6, E and F). The EP area showed a modest increase in FR irrespective of the envelope alteration. Whereas the mapping of the FR increases for the compressed meow generally is sparse and unclear, most of the neurons showing a decrease in FR to compressed or expanded meows are localized in the dorsal part of AI. This would suggest duration sensitivity, albeit not of the short-duration type (50–100 ms) found in the dorsal auditory cortex of the cat (He et al. 1997) for long-latency (>30 ms) neurons.

Windows of temporal integration

Although the importance of the temporal integration in complex sound processing is increasingly recognized, the features of such integration remain unclear, especially the length of a potential integration time window. Some neurons in AI respond to brief periodic stimuli only for repetition rates ≤20–40 Hz (Eggermont 2002; Lu et al. 2001; Schreiner et al. 1997). Reversing of local time segments in recorded speech does not affect its intelligibility if the segments are not >50 ms (Saberi and Perrott 1999). In the ferret, the mutual information between some calls and the response was found to reach a maximum when the temporal resolution of analysis was between 10 and 40 ms (Schnupp et al. 2006). From experiments on awake marmoset monkeys using periodic click trains, Wang et al. (2003) concluded that rapidly modulated signals would be integrated within a short-time window of about 20–30 ms. All these observations suggest that a temporal integration over a window between 10 and 50 ms long may occur when processing a more or less complex sound. However, the stronger response of a monkey to a call with two syllables compared with the response to each syllable alone implies the existence of an additional very long temporal integration window of several hundreds of milliseconds (Rauschecker and Tian 2000).

In fact, it is likely that several integration windows are involved in complex sound processing, from the shortest windows around 5 or 10 ms analyzing some aspects of roughness or rhythm, to the longest windows of several hundreds of milliseconds involved in message identifications like the association of two complex sounds or prosody. Using mutual information on increasing time intervals from stimulus onset between the stimuli (alterations of forward and time-reversed meows) and the FR associated with the responses as a quantification of discrimination ability for each site, our previous study showed that the inflection point in the cumulative increasing curve of mutual information occurred around 200 ms after meow onset (Gehr et al. 2000), making this earlier period crucial for discrimination of alterations. Our present results show that the strongest differences in FR between a meow and its altered versions or with silence generally occur during the same first 200 ms whatever the BF of the site (Fig. 9). The latency of the onset peaks (Fig. 8D) suggests that AI and perhaps PAF are the areas involved in this early processing. For longer times since onset the response to the meows is decreasing (Fig. 9B), except at some specific portions of stimuli, such as 750 ms for the natural meow. Furthermore, the differences in PSTH globally decrease with time since stimulus onset between the temporal altered forward stimuli in AI, but not in PAF and EP (Fig. 11). Meanwhile, neurons in EP reach their highest level of synchrony between 300 and 500 ms for a natural meow (Fig. 13). They also show the highest difference in FR between a natural meow and a temporally altered meow between 200 and 500 ms (Fig. 11). Two scales of temporal integration thus appear: 1) the first 200 ms are involved in identifying the early components of the stimulus, in AI and perhaps the Onset neurons of the PAF; 2) the remainder of stimulus is characterized by global inhibition in the Onset or FewPeaks neurons in AI, and by synchronized activity in EP areas, and probably encoding the features and altered features of the complex sound that constitute a vocalization.
Role of neural synchrony

Spontaneous synchrony is weak between distant sites (Fig. 12A) and between different cortical areas. Vocalization processing induces a general increase of synchrony between neurons compared with the spontaneous state (Fig. 12). The increase in synchrony is not caused by increased FR because the correction procedure used for the cross-correlation coefficient eliminates such effects (see METHODS). The increase in synchrony remains quite constant over stimulus time in AI, AAF, and PAF areas, but shows more temporal variability when sites in EP areas are involved (Fig. 13A). Except between EP and AI, the increase seems comparable between sites in different areas and within an area (Fig. 13). This emphasizes the existence of long-range intracortical connections. These results have to be put side by side with other studies on such connections: first, Eggermont showed that a stimulus such as a tone-pip or noise burst induced a higher increase in synchrony from a spontaneous condition (i.e., a post-stimulus condition) between neurons in different auditory areas than between neurons within the same area, mainly because the synchrony during silence is small between different areas (Figs. 4 and 5 in Eggermont 2000). Second, although there is evidence for tonotopic connections between auditory cortical areas (Lee et al., 2004a), distant, heterotopic connections between sites having a difference of >3 octaves between their characteristic frequency (CF) have also been well established (Lee et al. 2004b). Such connections may contribute to complex cortical processes also requiring horizontal intracortical connectivity (Kisvárday et al. 1996; Kubota et al. 1997; Read et al. 2001; Sutter et al. 1999). Neural synchrony seems to play a higher role in the processing of temporal changes than in carrier alterations (Fig. 14). Finally, we were often able to distinguish a forward from a reversed meow, even altered, by the mean synchrony values (Fig. 15C). The EP areas again seem to play an important role in this discrimination because the highest energy periods of the reverse and forward meows generate some high levels of synchrony within these areas (Fig. 13B).

Role of individual auditory areas

Vocalizations activate mainly AI and EP areas based on changes from spontaneous FR and these areas also show a high variability in these changes (Fig. 4C). The natural meow inhibited the response of neurons with sustained PSTH type in EP areas (Fig. 4A). The observed changes in neuronal firing suggest a specialization of some regions in the auditory cortex: the dorsal part of AI shows inhibition for carrier-altered and expanded vocalizations. The ventral and posterior parts of AI were strongly activated by the time-reversed meow. These results are difficult to relate to the general known properties of the parts of AI: the dorsal part of AI would be involved in analyzing complex patterns of frequency (Schreiner and Sutter 1992; Sutter and Schreiner 1991), whereas the ventral part would have a role in analyzing and detecting narrowband sounds (Schreiner and Sutter 1992; Sutter and Schreiner 1995). Nevertheless, we confirmed the distinct roles of the ventral and dorsal parts of AI in representing alterations of vocalizations. In general, the AAF does not appear very discriminating between the various vocalizations except for a very few neurons showing FR discrimination (Fig. 7A) or different temporal response compared with AI (Fig. 11C) between the forward and time-reversed meows.

The PAF has been shown to respond to tone onsets and to be as tonotopically organized as AI (Phillips and Orman 1984; Reale and Imig 1980). Compared with AI, PAF neurons exhibit longer response latencies, larger rebound responses, and lower temporal precision and are often tuned for intensity (Kitzes and Holligrigel 1996; Phillips and Orman 1984; Phillips et al. 1995, 1996). PAF is generally assumed to be involved in the analysis of acoustic signals of greater spectrotemporal complexity than AI (Loftus and Sutter 2001) and even implied in sound localization (Stecker and Middlebrooks 2003). PAF in cat (Heil and Irvine 1998) has been considered specialized for coding slowly varying sounds. The sensitivity of its neurons for rate and direction of FM sounds would make it suitable for detection and analysis of communication sounds (Tian and Rauschecker 1998). Our results show that the PAF was more sensitive to carrier alteration than to expansion or time-reversal of meows (Figs. 5, 6, and 11, B and C), suggesting a possible role in emitter identification. Its onset neurons also showed strong activation in response to a meow.

Little is known about the functional role of EP. Projections arise from the secondary auditory cortex and the insular and temporal cortical fields (Lee et al. 2004a) but also from AI, PAF, and visual nuclei (Winer et al. 2001). EP revealed interesting features for vocalization processing; the neurons with Sustained PSTH were inhibited by vocalizations, whereas the FewPeaks neurons were not. Moreover, most of the neurons in EP showed a higher FR for altered meows compared with natural meow. The synchrony found between sites in these areas also substantially varied during the time course of a natural meow or a time-reversed meow (Fig. 13). Finally, in EP the comparison of the responses between the natural meow and the time-reversed meow showed clear patterns of temporal differences, as well as those induced by the compression and the expansion of the natural meow (Fig. 11). EP therefore appears crucial for the interpretation of the information conveyed by the temporal alterations of a vocalization or by the interpretation of the relevance of some complex sounds such as time-reversed meows.

Which model for vocalization analysis?

Important preprocessing for complex sounds, such as vocalizations, is done subcortically, as reflected in the variability of PSTHs in response to vocalizations observed in the inferior colliculus of the guinea pig (Suta et al. 2003). Neurons in AI are sensitive to a large set of parameters characterizing simple sounds: frequency content, AM frequency, FM sweeps, and so forth. The call detectors hypothesis derives from these observations on simple sounds. It seems that if call-detecting neurons exist, they will be restricted to EP and ventral or posterior parts of AI. In addition, they may not be limited to the detection of one exact stimulus but more to a combination of features. In AI, the temporal processing of vocalizations is therefore of crucial importance. Several recent studies attempted to explain part of the temporal processing as a filtering of the stimulus by the spatiotemporal receptive field (STRF) or the tuning curve (deCharms et al. 1998; Depireux et al. 2001; Schnupp et al. 2001; Sen et al. 2001; Suta et al. 2003; Wang 2000; Wang et al. 1995). From the results presented here, it appears that this model is not sufficient for vocalizations processing by AI neurons for at least two reasons. First, whereas most peaks in the PSTH of neurons having a BF <5
kHz actually follow the rise of the temporal envelope of the stimulus filtered around the BF of the neuron (Fig. 8 in Gehr et al. 2000), numerous neurons with BFs above the frequency contents of the stimulus also generate onset and rebound responses (Fig. 10). Moreover, the filter model would predict the same PSTH shape and synchrony for compressed and expanded meows, which is definitely not the case (Figs. 11 and 14). Although the presence of call detectors in AI is very unlikely, the rationale for all the activity as generated by frequency filtering by the neuron remains unclear. These high-BF neurons might be sensitive to some temporally significant parameters as hypothesized by Steinmörch et al. (1990), who observed similar behaviors on neurons of monkeys responding to consonant–vowel syllables. More probably, however, their bandwidth is sufficiently high at 65-dB SPL and they may detect part of a strong temporal variation even at frequencies far below their BF. The ensuing spreading out of their onset and rebound response might also be a part of the code because relative latency was previously suggested as a possible code in auditory cortex (Eggermont 1998) and visual cortex (Gawne et al. 1996). The onset neurons of the PAF area may also help to identify some early more complex features.

After the onset and rebound generation, in the BF range \( \leq 5 \) kHz, the activity of responding neurons in AI is inhibited, whereas spectral components in that frequency range remain present in the stimulus (Fig. 9B). In the same time window the long-term integration of activity coming from AI is realized at least partially by the means of neural synchrony in EP (Fig. 13) and, perhaps, in PAF for spectral alterations. The information conveyed by the alteration of the natural meow could then be interpreted. The EP areas seem to integrate the information from the beginning of the stimulus presentation or at least as soon as the information has arrived in the area, about 30 ms later than in AI areas on average (Figs. 8D and 13A). This is in agreement with the hypothesis of Rauschecker and Tian (2000) that “auditory cortical pathways are organized in parallel as well as serially.” These authors previously showed that lateral belt areas of the superior temporal gyrus seem to be critically involved in the early processing of species-specific vocalizations as well as human speech. Because processing is contiguous in time in PAF and EP (Fig. 11), this does not seem compatible with the existence or the generalization of call detectors. As Schnupp wrote, “the role of AI in the processing of vocalization stimuli would be predominantly that of representing acoustic features of stimuli through temporal pattern codes in a nonselective manner, whereas a more auditory object-selective representation of the acoustic environment may well emerge in higher-order areas of auditory cortex” (Schnupp et al. 2006). The higher-order areas may not even be restricted to auditory areas because recently, in the primates, Romanski et al. (2005) recorded activation of the prefrontal cortical areas to vocalization stimuli and Gil-da-Costa et al. (2006) showed evidence with PET for higher-order processing of calls in the ventrolateral portions of the frontal cortex as well as the posterior parietal cortex and the posterior perisylvian cortex. Furthermore, as pointed out by Gehr et al. (2000), a rate-based representation of vocalization stimuli in cat AI would probably not be sufficient and certainly not efficient. We showed that PSTH types (temporal information), FR and neural synchrony allowed discriminating the features of altered and above all time-reversed vocalizations (Fig. 15). All these parameters and perhaps some others might be involved simultaneously in the processing and decoding of such complex sounds as vocalizations. The identification of the combinations of and interactions between neurons required to interpret vocalizations remains a challenge for future studies.

Acknowledgments

G. Shaw provided programming assistance. A. Noreia, M. Tomita, and N. Aizawa assisted with the data collection.

Grants

This work was supported by the Alberta Heritage Foundation for Medical Research, by the Natural Sciences and Engineering Research Council, and by the Campbell McLaurin Chair of Hearing Deficiencies.

References


