Differential Representation of Species-Specific Primate Vocalizations in the Auditory Cortices of Marmoset and Cat

XIAOQIN WANG AND SIDDHARTHA C. KADIA
Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received 26 March 2001; accepted in final form 25 June 2001

Wang, Xiaoqin and Siddhartha C. Kadia. Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 86: 2616–2620, 2001. A number of studies in various species have demonstrated that natural vocalizations generally produce stronger neural responses than do their time-reversed versions. The majority of neurons in the primary auditory cortex (A1) of marmoset monkeys responds more strongly to natural marmoset vocalizations than to the time-reversed vocalizations. However, it was unclear whether such differences in neural responses were simply due to the difference between the acoustic structures of natural and time-reversed vocalizations or whether they also resulted from the difference in behavioral relevance of both types of the stimuli. To address this issue, we have compared neural responses to natural and time-reversed marmoset twitter calls in A1 of cats with those obtained from A1 of marmosets using identical stimuli. It was found that the preference for natural marmoset twitter calls demonstrated in marmoset A1 was absent in cat A1. While both cortices responded approximately equally to time-reversed twitter calls, marmoset A1 responded much more strongly to natural twitter calls than did cat A1. This differential representation of marmoset vocalizations in two cortices suggests that experience-dependent and possibly species-specific mechanisms are involved in cortical processing of communication sounds.

INTRODUCTION

Species-specific vocalizations are communication sounds that many species rely on for their survival and social interactions. Communication sounds differ from other types of acoustic signals in that they are behaviorally relevant to a species. Although the biological importance of these acoustic signals is well recognized (Snowdon et al. 1982), their neural representation in the cerebral cortex has remained elusive (Wang 2000). A fundamental issue in understanding how the auditory system processes communication sounds is whether such sounds are processed differently from behaviorally irrelevant sounds. A long line of studies of cortical plasticity, both in development and adulthood (see reviews by Buonomano and Merzenich 1998; Schmidt et al. 1999), suggest that cortical representation of communication sounds should differ from that of behaviorally irrelevant sounds.

Preference of neural responses to natural vocalizations over time-reversed vocalizations has also been reported in various species such as bats (Esser et al. 1997), songbirds (Doupe and Konishi 1991; Margoliash 1983), and cats (Gehr et al. 2000). Wang et al. (1995a) studied responses to both natural and synthetic vocalizations in populations of neurons in the primary auditory cortex (A1) of a highly vocal primate, the common marmoset (*Callithrix jacchus jacchus*). It was found that the majority of neurons showed stronger responses to natural marmoset twitter calls than to time-reversed twitter calls. Time-reversed calls do not bear the behavioral meaning associated with natural twitter calls but have the same spectral contents and similar acoustic complexity as the natural calls. While this finding has implications for the role of behavioral relevance and species-specificity underlying cortical responses, it alone does not rule out the possibility that differences in cortical responses may simply be due to differences in the acoustic structures between the natural and time-reversed calls. One way to address this issue is to study neural responses to marmoset twitter calls, in both natural and time-reversed forms, in the auditory cortex of another species that does not encounter marmoset vocalizations in its acoustic environment. We have performed this comparative analysis in domestic cats, a mammalian species whose A1 shares a number of similar anatomical and physiological properties with A1 of primates. We reasoned that the differences between responses of cat A1 to natural and time-reversed marmoset twitter calls should solely be due to differences in the acoustic structures of these two types of sounds because neither bears any behavioral relevance to the cats used in our study. The results of this study showed that neurons in cat A1 did not exhibit preference to marmoset natural twitter calls as observed in marmoset A1 and therefore suggest experience-dependent and possibly species-specific mechanisms underlying cortical processing of behaviorally relevant vocalizations in the auditory cortex of marmosets.

METHODS

Animal preparation and recording procedures

Surgical procedures were described in details in a previous study (Lu and Wang 2000). Anesthesia was maintained throughout an experiment by intravenous injection of pentobarbital sodium. The location of A1 was confirmed electrophysiologically. Recording experiments were conducted with the animal placed within a double-wall soundproof chamber (IAC-1024) whose interior was covered by 3-in acoustic absorption foam (Sonex, Illbruck). Multi-unit extracel-

Address for reprint requests: X. Wang, Dept. of Biomedical Engineering, Johns Hopkins University School of Medicine, 720 Rutland Ave., Ross 424, Baltimore, MD 21205 (E-mail: xwang@bme.jhu.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

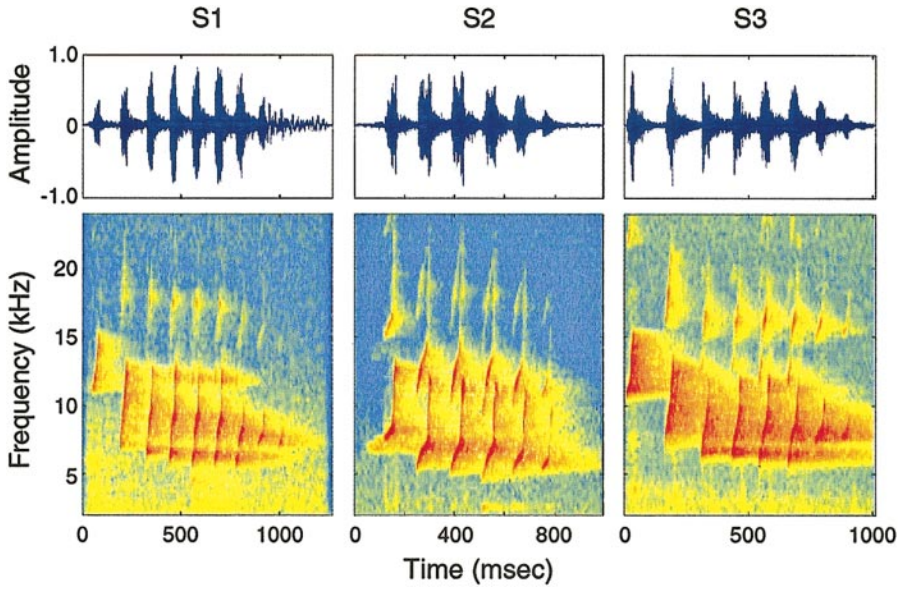

FIG. 1. Natural marmoset twitter calls used as stimuli in both marmoset and cat experiments. Each of the 3 twitter calls (S1–S3) was recorded from an individual marmoset monkey. The waveform (*top*) and spectrogram (*bottom*) of each call are shown. The amplitude of the waveform was normalized to within (−1.0, 1.0) for display purpose. The frequency range shown in the spectrograms is 2–24 kHz. For each natural twitter call, a time-reversed call was generated for the experiments (not shown).

lular recordings were made using tungsten microelectrodes (Microprobe, 1–2 MΩ impedance at 1 kHz) from the middle cortical layers (depths of 600–900 μm). The characteristic frequency (CF) and threshold were identified using a manually controlled oscillator and attenuator. Neural activity was amplified and filtered at 0.3–7 kHz. Action potentials were detected by a window discriminator. A1 was systematically sampled from dorsal to ventral regions. These experimental conditions and recordings procedures were comparable to those used in earlier marmoset experiments (Wang et al. 1995a) from which the comparison data were obtained.

Vocalization stimuli

Three pairs of natural and time-reversed twitter calls (identical to those used in earlier marmoset experiments) were studied in the cat experiments. Figure 1 displays the three natural twitter calls. Details on acquisition of these vocalizations were given in Wang et al. (1995a). A time-reversed twitter call (hereafter referred to as *reversed call*) was generated by reversing the time course of a natural twitter call. Acoustic stimuli were delivered under free-field conditions by a speaker located ~1 m in front of the animal. The speaker (XTS-35, Radio Shack) had a flat (±5 dB) frequency response from 100 Hz to 20 kHz. Vocalization stimuli were generated through a 16-bit D/A converter at 48-kHz sampling rate and delivered at 60–70 dB SPL. All stimuli, 20 repetitions each, were presented in random order. Inter-stimulus intervals were >1 s.

Data analysis

The results presented in this report were based on 70 units recorded from A1 of two cats with CFs ranging from 4 to 9 kHz. Responses of marmoset A1 neurons were based on Wang et al. (1995a) and subsequent electrophysiological experiments performed under identical conditions. Eighty-nine units recorded from A1 of two marmosets with CFs between 4 and 9 kHz were analyzed for direct comparison with the data obtained from the cat experiments. Responses of these units to the three pairs of natural and reversed calls as used in the cat experiments were analyzed for the present study. The multi-unit responses from marmoset A1 were recorded with the same brand of electrode (Microprobe, 1–2 MΩ), at similar recording depths (middle cortical layers) and under the same anesthetic condition (barbiturate anesthesia) as in the cat experiments. The vocalization stimuli were delivered at 65 dB SPL, 10–20 repetitions per stimulus.

Mean firing rate, calculated over the stimulus duration, was used to measure responses to natural and reversed calls for both marmoset and cat data. We chose to use this simple measure because it reflects the overall responsiveness of the neurons studied in both species and does not involve other assumptions. Spontaneous discharge rate was estimated from recording intervals prior to stimulus presentations and removed from the mean firing rate. A *selectivity index* (*d*) was used to quantify the difference between responses to a pair of natural and reversed calls on a unit by unit basis and is defined as follows

$$d = (R_{\mathrm{Nat}} - R_{\mathrm{Rev}})/(R_{\mathrm{Nat}} + R_{\mathrm{Rev}})$$

Equivalently

$$R_{\mathrm{Nat}}/R_{\mathrm{Rev}} = (1 + d)/(1 - d)$$

where R_{Nat} and R_{Rev} are mean firing rates due to the natural and reversed calls, respectively. A *d* value of 1.0 (−1.0) indicates that a neuron responded only to the natural (reversed) call. For each stimulus pair, only the units with the minimum mean firing rate (either R_{Nat} or R_{Rev}) > 3 spikes/s were included in the analyses. Statistical significance between response measures was evaluated using a *t*-test. $P < 0.001$ is considered statistically significant (Tables 1 and 2).

TABLE 1. *Comparison of three response measures between marmosets and cats*

	d				R_{Nat}				R_{Rev}			
Stimulus	S1	S2	S3	All	S1	S2	S3	All	S1	S2	S3	All
Marmoset	0.479 ± 0.361 (89)	0.335 ± 0.302 (76)	0.385 ± 0.340 (89)	0.403 ± 0.341 (254)	11.87 ± 7.18 (89)	11.70 ± 8.54 (76)	12.6 ± 7.62 (89)	12.08 ± 7.74 (254)	5.61 ± 7.11 (89)	6.49 ± 5.52 (76)	6.88 ± 7.52 (89)	6.32 ± 6.82 (254)
Cat	0.047 ± 0.265 (67)	0.086 ± 0.238 (68)	0.068 ± 0.372 (70)	0.068 ± 0.297 (205)	6.05 ± 2.73 (67)	6.31 ± 3.60 (68)	5.54 ± 3.75 (70)	5.96 ± 3.39 (205)	5.66 ± 2.99 (67)	5.38 ± 3.35 (68)	4.78 ± 3.05 (70)	5.27 ± 3.14 (205)
P	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.952	0.152	0.029	0.042

Numbers given are means ± SD (number of samples in parentheses). *P* is unpaired *t*-test score comparing the marmoset versus the cat.

TABLE 2. *Statistical significance test between R_{Nat} and R_{Rev}*

	Marmoset			Cat		
Stimulus	S1	S2	S3	S1	S2	S3
P	<0.0001	<0.0001	<0.0001	0.272	0.010	0.047

P is paired *t*-test score comparing R_{Nat} versus R_{Rev}. Numerical measures of R_{Nat} and R_{Rev} (means ± SD) are listed in Table 1.

RESULTS

Natural twitter calls are composed of a series of "phrases," each of which is made of several upward FM sweeps and their harmonics (Fig. 1). The spectral components of the first harmonic are centered near 7 kHz and spread to as low as 4–5 kHz in samples across different animals (Agamaite and Wang 1997; Epple 1968). A reversed call has spectral contents occupying the same frequency range as does the natural call but has downward FM sweeps instead and reversed time courses of their amplitudes (as reflected in its envelope). The overall energy of a reversed call is identical to that of a natural twitter call.

In Fig. 2*A*, mean firing rates of responses to a pair of natural and reversed calls are compared for the neurons studied in cat A1. Overall, responses of 69 units are distributed around the dashed line that indicates equal discharge rates to both stimuli. The difference between responses to the two stimuli is further quantified by a selectivity index (see METHODS). The distribution of the selectivity index (Fig. 2*B*) has a mean of 0.047

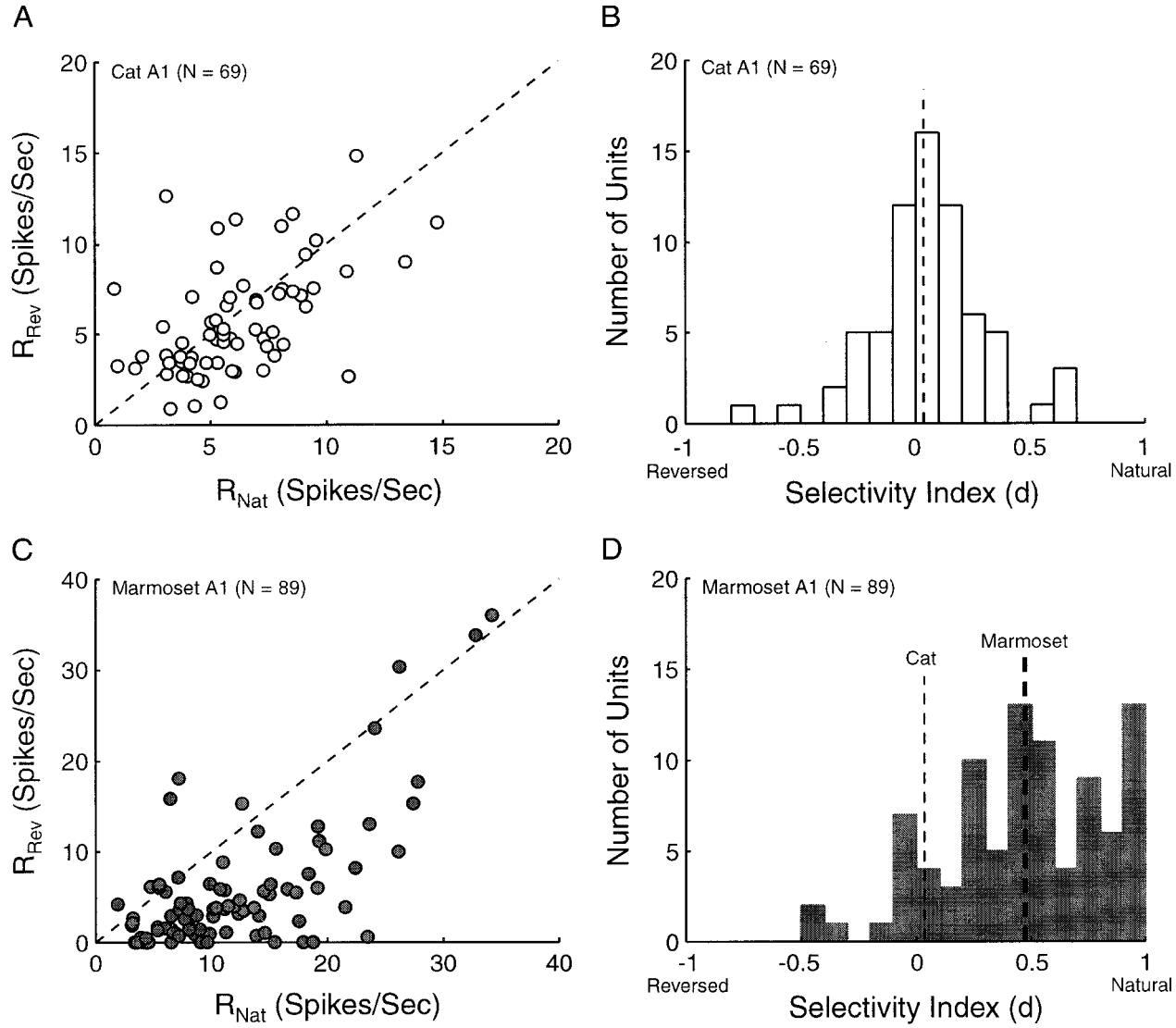

FIG. 2. Cortical responses to a representative pair of natural and reversed twitter calls (stimulus: S1, Fig. 1, *left*). *A*: comparison between mean firing rates to natural and reversed calls (R_{Nat} and R_{Rev}) for individual cortical units studied in cats. - - -, equal responses to the natural and reversed calls. *B*: histogram showing the distribution of the selectivity index (*d*) for cat A1 units (binwidth, 0.1). The mean of the distribution (0.047 ± 0.265, Table 1) is marked (- - -); the equivalent R_{Nat}/R_{Rev} ratio is 1.1 (see METHODS). *C*: comparison between R_{Rev} and R_{Nat} for individual cortical units studied in marmosets. - - -, equal responses to the natural and reversed calls. *D*: distribution of the selectivity index for marmoset A1 units (binwidth, 0.1) with the mean (0.479 ± 0.361, Table 1) marked by the thick - - - (*right*); the equivalent R_{Nat}/R_{Rev} ratio is 2.8. The mean of the selectivity index for cats (from *B*) is also shown for comparison (- - -, *left*). The selectivity index for 2 species differs significantly in their means ($P < 0.0001$, Table 1).

(equivalent to a 1.1 ratio of R_{Nat}/R_{Rev}), indicating that natural and reversed calls produced approximately equal responses over the population of the neurons studied. This is in contrast to the response characteristics of neurons in marmoset A1 where the same natural twitter call resulted in greater mean firing rates than did the corresponding reversed call in most of the 89 units studied (Fig. 2*C*). The distribution of the selectivity index of marmoset A1 units (Fig. 2*D*) has a mean of 0.479 (equivalent to a 2.8 ratio of R_{Nat}/R_{Rev}), which is far greater than that of cat A1 neurons. The distributions of the selectivity indices of two species are statistically different ($P < 0.0001$, Table 1).

Quantitative analyses of responses to all three pairs of natural and reversed calls are shown in Fig. 3. As is seen in Fig. 2, the disparity in neural selectivity between cat A1 and marmoset A1 is evident in all three cases. In each case, neurons in marmoset A1 showed significantly higher selectivity index than did neurons in cat A1 (Fig. 3*A*, Table 1). The average selectivity index over three cases is 0.068 in cats and 0.403 in marmosets ($P < 0.0001$, Table 1). The greater selectivity index in marmoset A1 resulted from stronger responses to the natural twitter calls (Fig. 3*B*). On average, mean firing rate of marmoset A1 neurons is about twice as high as that of cat A1 neurons in response to natural twitter calls (12.08 vs. 5.96 spikes/s, $P < 0.0001$, Table 1). Responses to reversed calls (Fig. 3*C*), on the other hand, have similar magnitudes in both species. Mean firing rate to reversed calls averaged over three cases shows a small but nonsignificant difference between marmosets and cats (6.32 vs. 5.27 spikes/s, $P = 0.042$, Table 1). Furthermore, responses to natural twitter calls are significantly larger than responses to reversed calls in marmosets for all stimulus pairs tested (Tables 1 and 2). In cats, only a small difference is observed in the same comparison (Tables 1 and 2). In summary, these results clearly showed that, on average, neurons in marmoset A1 responded more strongly to natural twitter calls than to their time-reversed versions, whereas neurons in cat A1 responded to both types of sounds approximately equally.

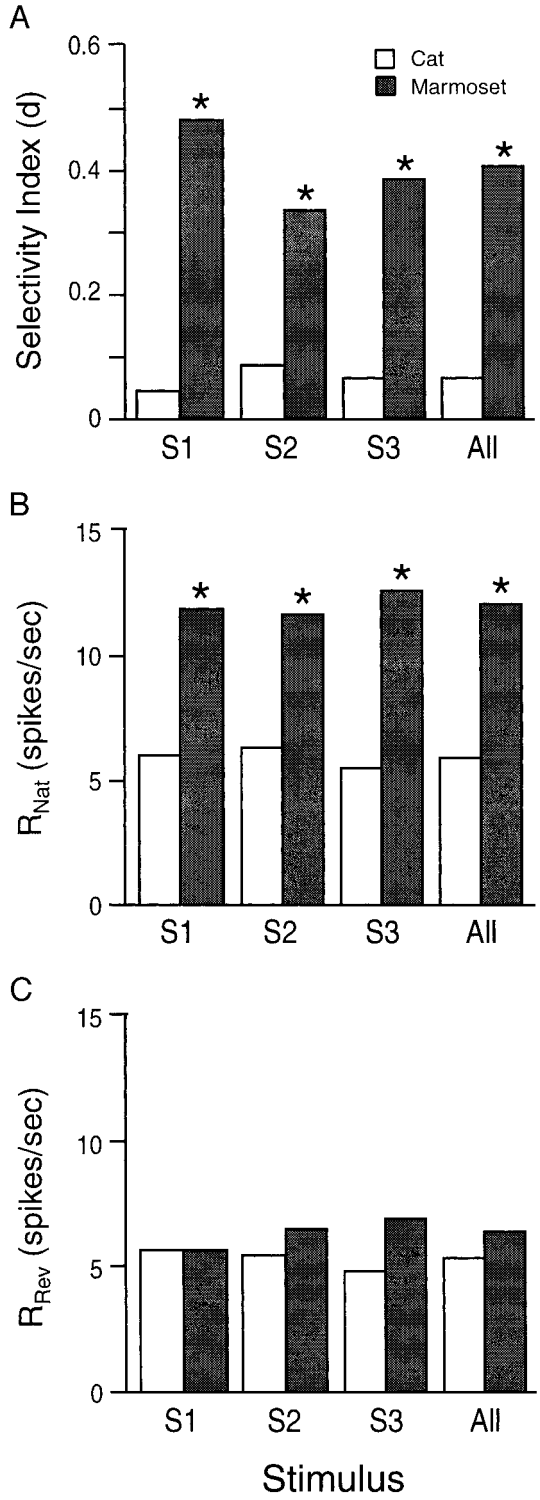

FIG. 3. Average selectivity index and response magnitudes are compared between cats (□) and marmosets (■) for 3 pairs of twitter calls (S1, S2 and S3). The measures averaged over all stimulus conditions are also shown (All). *A*: selectivity index (*d*). *B*: mean firing rate of responses to natural calls (R_{Nat}). *C*: mean firing rate of responses to reversed calls (R_{Rev}). The numerical values are listed in Table 1. *, statistical significance at $P < 0.001$ (*t*-test) for comparisons between marmosets and cats under each stimulus condition (Table 1).

DISCUSSION

The A1 of cats shares some similarities with the A1 of primates, including marmosets. In both species, A1 receives thalamic inputs predominantly from the ventral division of the medial geniculate body (Aitkin and Park 1993; Middlebrooks and Zook 1983; Morel and Kaas 1992; Winer 1992). A tonotopic map exists in A1 of both species (Aitkin et al. 1986; Imig et al. 1977; Merzenich and Brugge 1973; Merzenich et al. 1975). Although the CF progresses from low frequency to high frequency rostrocaudally in marmoset A1, whereas it progresses in the opposite direction in cat A1, there have been no reports of major differences between the two species in basic functional properties of A1 as determined by pure tone responses. The hearing ranges of marmosets and cats overlap at ~1–20 kHz (Fay 1988). The frequency range of marmoset twitter calls is well within sensitive hearing frequencies of cats. The similar response magnitudes for the reversed calls between marmoset and cat (Fig. 3*C*) are an assuring indication that auditory cortices of both species are capable of responding to sounds with the acoustic complexity such as that found in marmoset twitter calls. It further indicates that there was no systematic bias in these two sets of data obtained from separate experiments.

The natural twitter calls used in this study are biologically

important to marmosets. Twitters are social calls in the marmoset's vocal repertoire (Epple 1975) and are frequently used in vocal exchanges between members of a colony (Wang 2000). They bear, however, no behavioral relevance to the cats under study, which were never exposed to marmoset vocalizations. The time-reversed twitter calls are not behaviorally meaningful to both marmosets and cats. The differential responses due to natural and reversed calls observed in marmosets, but absent in cats, suggest that neural mechanisms other than those responsible for encoding acoustic structures of complex sounds are involved in cortical processing of behaviorally important communication sounds. These mechanisms may include experience-based and developmental plasticity (functionally and/or structurally) as well as predisposed and specialized circuitry in the marmoset's auditory system that does not exist in cats. Such mechanisms represent the adaptation of the auditory system to sensory environment over different time scales (from days, years, to generations). It has been demonstrated that learning spatial-temporal sensory input patterns resulted in reorganization of response properties accordingly in the sensory cortices of primates (Recanzone et al. 1992; Wang et al. 1995b). There was also evidence in songbirds that neurons selective to bird's own song emerged from development (Doupe 1997). A recent study showed that plasticity of the cochleotopic map has different characteristics in the auditory cortices of mustached bats and Mongolian gerbils (Sakai and Suga 2001), suggesting species-specific mechanisms operating at A1. While it is unclear which or all of these mechanisms have contributed to the observed differential responses, the findings of this study suggest that experience-dependent and possibly species-specific mechanisms are involved in cortical processing of behaviorally important communication sounds. It remains unclear whether the observed disparity in cortical responses was created in A1 or might have been contributed by subcortical processing. Our result has direct implications for interpreting neural responses to natural and artificial sounds. It argues strongly that the exploration of cortical mechanisms responsible for encoding communication sounds must be based on biologically meaningful models.

We thank Drs. Michael Merzenich and Christoph Schreiner for supporting the marmoset experiments referred to in this study as well as Drs. Ralph Beitel and Steven Cheung, who participated in those experiments. We thank Dr. Edward Bartlett and T. Lu for helpful comments on the manuscript and A. Pistorio for proofreading the manuscript.

This research was supported by National Institute on Deafness and Other Communication Disorders Grant DC-03180 and by a Presidential Early Career Award for Scientists and Engineers (X. Wang).

REFERENCES

AGAMAITE JA AND WANG X. Quantitative classification of the vocal repertoire of the common marmoset (*Callithrix jacchus jacchus*). *Assoc Res Otolaryngol Abstr* 20: 144, 1997.

AITKIN LM, MERZENICH MM, IRVINE DR, CLAREY JC, AND NELSON JE. Frequency representation in auditory cortex of the common marmoset (*Callithrix jacchus jacchus*). *J Comp Neurol* 252: 175–185, 1986.

AITKIN L AND PARK V. Audition and the auditory pathway of a vocal new world primate, the common marmoset. *Prog Neurobiol* 41: 345–367, 1993.

BUONOMANO DV AND MERZENICH MM. Cortical plasticity: from synapses to maps. *Annu Rev Neurosci* 21: 149–186, 1998.

DOUPE AJ. Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J Neurosci* 17: 1147–1167, 1997.

DOUPE AJ AND KONISHI M. Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad Sci USA* 88: 11339–11343, 1991.

EPPLE G. Comparative studies on vocalization in marmoset monkeys (*Hapalidae*). *Folia Primatol* 8: 1–40, 1968.

EPPLE G. The behavior of marmoset monkeys (Callithricidae). In: *Primate Behavior: Developements in Field and Laboratory Research,* edited by Rosenblum LA. New York: Academic, 1975, vol. 4, p. 195–239.

ESSER KH, CONDON CJ, SUGA N, AND KANWAL JS. Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat Pteronotus parnellii. *Proc Natl Acad Sci USA* 94: 14019–14024, 1997.

FAY RR. *Hearing in Vertebrates: A Psychophysics Databook*. Winnetka, IL: Hill-Fay Associates, 1988.

GEHR DD, KOMIYA H, AND EGGERMONT JJ. Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. *Hear Res* 150: 27–42, 2000.

IMIG TJ, RUGGERO MA, KITZES LM, JAVEL E, AND BRUGGE JF. Organization of auditory cortex in the owl monkey (Aotus trivirgatus). *J Comp Neurol* 177: 111–128, 1977.

LU T AND WANG X. Temporal discharge patterns evoked by rapid sequences of wide- and narrowband clicks in the primary auditory cortex of cat. *J Neurophysiol* 84: 236–246, 2000.

MARGOLIASH D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J Neurosci* 3: 1039–1057, 1983.

MERZENICH MM AND BRUGGE JF. Representation of cochlear partition on the superior temporal plane of the macaque monkey. *Brain Res* 50: 275–296, 1973.

MERZENICH MM, KNIGHT PL, AND ROTH G. Representation of cochlea within primary auditory cortex in the cat. *J Neurophysiol* 38: 231–249, 1975.

MIDDLEBROOKS JC AND ZOOK JM. Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. *J Neurosci* 3: 203–224, 1983.

MOREL A AND KAAS JH. Subdivisions and connections of auditory cortex in owl monkeys. *J Comp Neurol* 318: 27–63, 1992.

RECANZONE GH, SCHREINER CE, AND MERZENICH MM. Plasticity in the primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci* 13: 87–104, 1992.

SAKAI M AND SUGA N. Plasticity of the cochleotopic (frequency) map in specialized and nonspecialized auditory cortices. *Proc Natl Acad Sci USA* 98: 3507–3512, 2001.

SCHMIDT KE, GALUSKE RA, AND SINGER W. Matching the modules: cortical maps and long-range intrinsic connections in visual cortex during development. *J Neurobiol* 41: 10–17, 1999.

SNOWDON CT, BROWN CH, AND PETERSEN MR. *Primate Communication.* Cambridge, UK: Cambridge Univ. Press, 1982.

WANG X. On cortical coding of vocal communication sounds in primates. *Proc Natl Acad Sci USA* 97: 11843–11849, 2000.

WANG X, MERZENICH MM, BEITEL R, AND SCHREINER CE. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74: 2685–2706, 1995a.

WANG X, MERZENICH MM, SAMESHIMA K, AND JENKINS WM. Remodeling of hand representation in adult cortex determined by timing of tactile stimulation. *Nature* 378: 71–75, 1995b.

WINER JA. The functional architecture of the medial geniculate body and the primary auditory cortex. In: *The Mammalian Auditory Pathway: Neuroanatomy,* edited by Popper N and Fay RR. New York: Springer-Verlag, 1992, p. 222–409.