J Comp Physiol A (2006) 192: 27–38
DOI 10.1007/s00359-005-0043-z

ORIGINAL PAPER

Cory T. Miller · Xiaoqin Wang

Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets

Received: 8 February 2005 / Revised: 11 July 2005 / Accepted: 11 July 2005 / Published online: 31 August 2005

Abstract A fundamental issue in neuroscience pertains to how different cortical systems interact to generate behavior. One of the most direct ways to address this issue is to investigate how sensory information is encoded and used to produce a motor response. Antiphonal calling is a natural vocal behavior that involves individuals producing their species-specific long distance vocalization in response to hearing the same call and engages both the auditory and motor systems, as well as the cognitive neural systems involved in decision making and categorization. Here we present results from a series of behavioral experiments investigating the auditory–vocal interactions during antiphonal calling in the common marmoset (*Callithrix jacchus*). We manipulated sensory input by placing subjects in different social contexts and found that the auditory input had a significant effect on call timing and propensity to call. Playback experiments tested the significance of the timing of vocal production in antiphonal calling and showed that a short latency between antiphonal calls was necessary to maintain reciprocal vocal interactions. Overall, this study shows that sensory-motor interactions can be experimentally induced and manipulated in a natural primate vocal behavior. Antiphonal calling represents a promising model system to examine these issues in non-human primates at both the behavioral and neural levels.

Keywords Antiphonal calling · Common marmoset · Neuroethology · Animal communication · Sensory-motor

C. T. Miller (✉) · X. Wang
Laboratory of Auditory Neurophysiology,
Department of Biomedical Engineering,
Johns Hopkins University School of Medicine,
720 Rutland Ave, Traylor 412,
Baltimore, MD, 21205 USA
E-mail: ctmiller@bme.jhu.edu
Tel.: +1-410-5026968
Fax: +1-410-6149599

Abbreviations CM: Cagemate · HVC: High vocal center · NCM: Non-cagemate · NCM-OS: Non-cagemate of the opposite sex · NCM-SS: Non-cagemate of the same sex

Introduction

Vocal communication is a dynamic process involving an interaction of several neural systems (Nottebohm et al. 1976; Liberman 1996; Ghazanfar and Hauser 2001). At its core, the auditory and motor systems are inextricably linked to vocal behavior because vocalizations are both perceived and produced. These systems do not, however, function in isolation. But rather represent different components of an integrated system that complement each other in order to generate natural vocal behaviors. While the auditory and motor systems are traditionally studied independently, recent evidence suggests that sensory-motor interactions modulate the properties of neural substrates known to be key to vocal communication (Jarvis et al. 1998; Hessler and Doupe 2000; Eliades and Wang 2003).

Songbirds have emerged as the most well-studied neural system in animal vocal communication (Hauser 1996; Carew 2000; Marler and Slabbekoorn 2004). Traditionally, much of this research has been focused on the representation of song in either the sensory (Sen et al. 2001; Grace et al. 2003) or motor (Yu and Margoliash 1996; Spiro et al. 1999) system. More recently, however, data indicate that these representations are not static, but rather modulate due to interactions between the sensory and motor systems (Dave et al. 1998; Schmidt and Konishi 1998). For example, the properties of HVC neurons change depending on whether the male is producing song directed at a female or freely singing, despite the fact that there is no difference in the acoustic structure of the song produced in these two contexts (Jarvis et al. 1998; Hessler and Doupe 2000). These results suggest that if we are to understand the neural

mechanisms underlying vocal communication, more detailed inquiries into sensory-motor interactions are needed.

Similarly to many taxonomic groups, vocal communication plays a central role in primate behavior. Unlike songbirds, however, there are few data examining the mechanisms underlying vocal production and perception in primates. This dearth of information is in part due to a lack of a stereotyped vocal behavior in primates analogous to other neuroethological systems, such as birdsong, that can be experimentally induced, manipulated, and repeated under laboratory conditions. One exception is antiphonal calling (Ghazanfar et al. 2001, 2002; Miller et al. 2001a, 2005; Miller and Hauser 2004). Antiphonal calling occurs when one animal produces a long distance contact call and a second visually occluded animal emits the same type of vocalization in response. This reciprocal vocalization is termed the 'antiphonal call' (Fig. 1a). Because antiphonal calling involves animals both hearing and producing vocalizations, it represents a natural (i.e., untrained), vocal behavior that is ideal for explorations into the behavioral and neural processes involved in sensory-motor interactions, a relationship well suited for neuroethological inquiries of primate vocal behavior (Ghazanfar and Hauser 2001).

Here we present experiments that examine auditory–vocal interactions during antiphonal calling in the common marmoset (*C. jacchus*). During antiphonal calling, common marmosets produce the 'phee' (Fig. 1a, b), their species-specific contact call (Norcross and Newman 1993, 1997; Norcross et al. 1994). A wealth of data already exist on the neural representation of acoustic signals (Lu et al. 2001a, b; Barbour and Wang 2003a, b), including vocalizations (Wang et al. 1995; Wang 2000; Wang and Kadia 2001), in the common marmoset auditory cortex. A well-defined vocal behavior in laboratory conditions, such as antiphonal calling, is essential to pursue neurophysiological studies of the brain circuitry underlying auditory–vocal interactions in nonhuman primates. This study represents the first step towards developing the antiphonal calling behavior into a neuroethological model of vocal communication.

In this paper, we present two experiments designed to test how changes in sensory input affect the motor output in antiphonal calling. The goal of the first set of experiments was to characterize the marmosets' natural vocal behavior and develop a working definition of antiphonal calling in this species. To this end we recorded the long-distance vocal behavior of marmosets in four different social contexts. Although subjects were visually occluded during these experiments, information about the caller's identity and sex is encoded in the vocalizations of marmosets and other closely related species (Norcross and Newman 1993; Miller et al. 2001b, 2004) and therefore could be used to modulate aspects of the vocal behavior. Building on these results in Experiment 2, we manipulated the timing of stimulus presentation in a series of playback experiments designed to test the significance of the temporal intervals between vocalizations during antiphonal calling. For these experiments, we used the same vocalizations, hence controlling for the information content of the signal, but we presented the stimuli in either an interactive or static manner. Overall, we predicted that changes in the sensory input would predictably modulate the subjects' volubility and the timing of vocal production.

Methods

Subjects

Seven adult common marmosets (four male/three female) housed at the Johns Hopkins University School of Medicine served as subjects in this study. The same subjects participated in both Experiment 1 and 2. Three of the males and two of the females were adult pair-bonded individuals from different cages in colony. The third female was an adult offspring from the cage of one of the adult male subjects. The fourth male in the study was pair-bonded with an adult female, but this female was not included in the study because of difficulties in familiarizing her to leave the cage in a transport box. Although she was used in the CM condition of the first experiment, her data were not analyzed. All of these adult pairs had 2–4 juvenile offspring living in the colony cage with them. As the experiments were conducted over 10 months, the exact number of offspring was different at various points of the study. Subjects' diet consisted of a combination of monkey chow, fruit, and yogurt. All animals had ad libitum access to water. We conducted all experiments during daylight hours between 0800 and 1800.

Effects of social context (experiment 1)

Aim

The goal of this experiment was twofold. We first aimed to characterize the temporal patterning of phee production during the natural long distance vocal behavior of common marmosets. And second, to use these data to determine what types of vocal behavior could be classified as antiphonal. This was accomplished by recording subjects' natural vocal behavior in four contexts. We predicted that social context (i.e., sensory input) would have a significant effect on subjects' vocal behavior.

Procedure

We transported subjects from the colony to the testing room in transport cages. During transport, we precluded any visual recognition of the other individual in the experiment by insuring that the subjects were visually

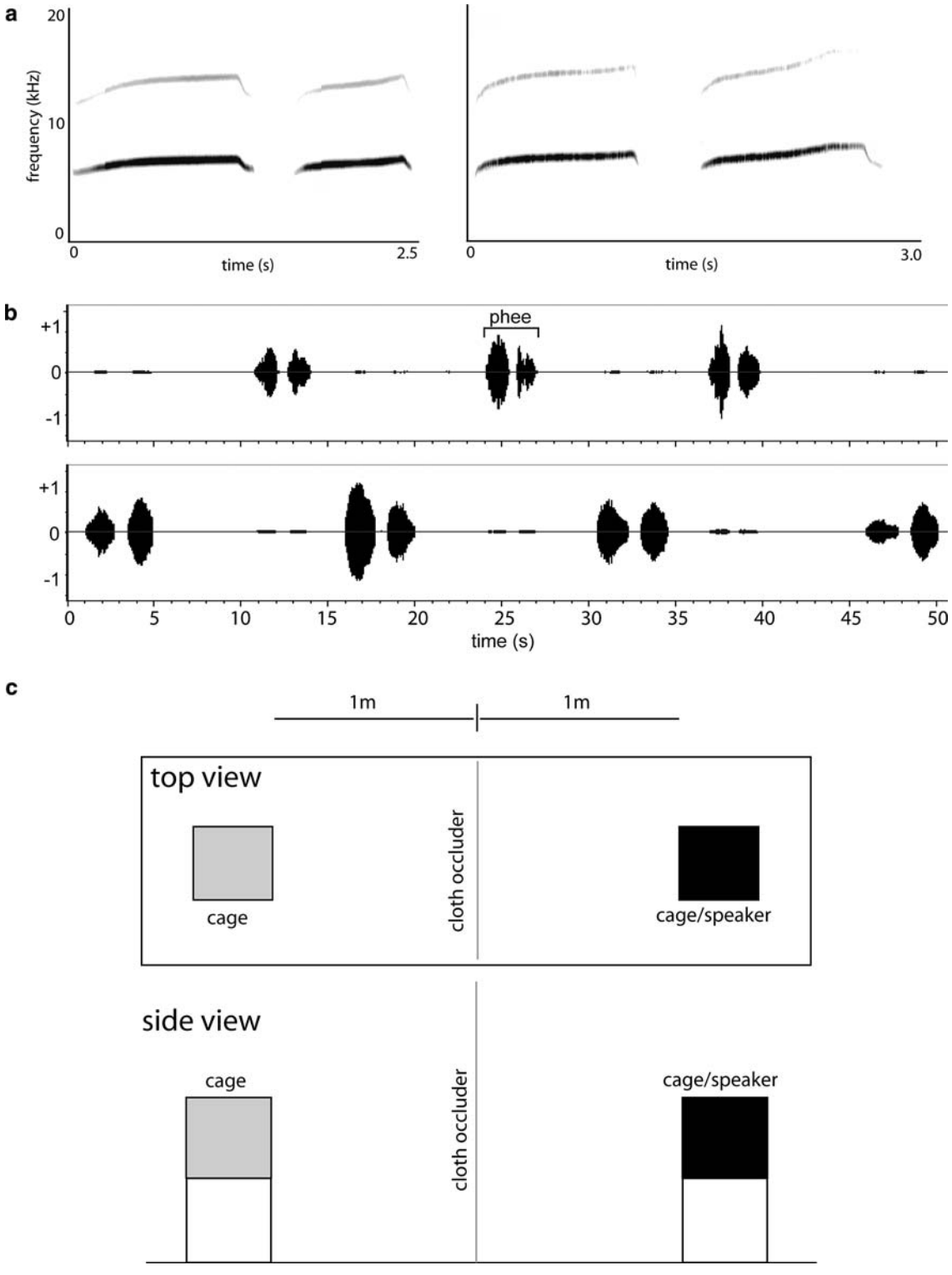

Fig. 1 Call type and experimental setup. **a** A spectrogram (frequency: *y*-axis; time: *x*-axis) of 'phee' vocalizations produced by two different individuals. **b** Time-amplitude waveforms (*y*-axis: dB; *x*-axis: time) of two individual marmosets engaged in a consecutive antiphonal calling exchange. The waveform of the 'phee' vocalization is marked. **c** Schematic drawing of the experimental setup with the top (above) and bottom (below) views shown. The experimental cage to the left is separated from the cage/speaker by 2 m with a cloth occluder placed equidistant between the two

occluded from each other at all times. The testing room was 7 m×4 m in size and had the walls covered completely in acoustic attenuating foam and a carpet floor. Once inside the room, we placed the subjects in wire mesh cages—each animal in an individual cage—separated by 2 m with an opaque cloth occluder equidistant between the two cages (Fig. 1c). As such the animals could interact vocally, but could not obtain visual cues

from each other during the length of the experiment. We aimed a Sennnheiser directional microphone (ME 66) at each cage and recorded all natural vocal behavior directly to the hard drive on either an Apple G4 powerbook or G5 Desktop computer using a Digidesign Mbox I/O device. Each test session lasted for 15 min. Following an experiment, we returned subjects to their home cage and cleaned the cages in the test room. As a final step in cleaning, we sprayed the cages with kennel odor eater (Thornell Corporation, Smithville, MO, USA)—a product which effectively masks potential olfactory cues during testing. Given the use of this product, the distance between subjects during testing and the short length of the experiments, it is doubtful that sufficient olfactory cues were available to subjects during testing to recognize the other individual in the experiment with these cues. We cannot, however, completely rule out the possibility that such cues could have been used at times.

We ran all subjects on four different conditions in this experiment. Three of these conditions consisted of pairing animals with individuals of different social categories: cagemate (CM), non-cagemate of the same sex (NCM-SS) and non-cagemate of the opposite sex (NCM-OS). The fourth condition involved placing a single animal in the test cage (ALONE) and recording their vocal behavior for the same length of time. In the CM condition, subjects were always paired with their mate except in one pair due to consistent difficulty removing the mate from the home cage in the colony. We ran subjects on each of these conditions three times in a randomized order. No subject participated in more than one test session a day.

Analysis

We analyzed the vocal behavior during each test session using Raven software (Cornell Bioacoustics Research Program). Approximately 95% of the vocalizations produced in our experiment were phees. As such, all analyses focused only on this call type. The duration of time between the offset of each 'phee' vocalization and the onset of the subsequent 'phee' vocalization was measured. We grouped 'phee' calls into two categories based on the preceding vocal behavior: 'within-individual' and 'between-individual'. Any phee that had the preceding phee produced by the same individual was categorized as a 'within-individual' call. During instances when the preceding phee was produced by a different individual, we classified the phee as a 'between-individual' call. As the ALONE trials consisted of only one subject, all calls produced in this context are deemed 'within-individual'. Our rationale for using these working classifications was the following. Antiphonal calls are by definition vocal responses to a conspecific vocalization that happen over short time scale. The temporal interval over which antiphonal calls occur. But may vary between species. We felt that distinguishing between phees produced following a phee by a conspecific from those following a call produced by the same individual without imposing an a priori assumption about the time scale over which the vocal behaviors occurred was an appropriate way to ascertain at least a working definition of antiphonal calling in marmosets. We leave open the possibility that continued study will lead to a refinement of the resulting definition. All statistical analyses of data on subjects' latency to call were performed using unpaired two-tailed *t*-tests. A repeated-measures ANOVA was used to compare the number of calls produced in these experiments.

Experimentally induced antiphonal calling (experiment 2)

Here we sought to replicate marmoset antiphonal calling behavior in a playback experiment by testing the significance of the reciprocal nature of this vocal behavior. We conducted three experimental conditions in which we manipulated either the stimulus type or temporal pattern of stimulus presentation. We predicted that temporally synchronized antiphonal calling is necessary to maintain natural levels of antiphonal calling.

Procedure

Following the procedure in Experiment 1, we transported subjects into the same testing room via a small transport box. The only difference from the previous experiment is that one of the wire mesh cages was replaced with a speaker. Once the subjects were situated in the testing cage, we began a playback experiment. For Experiment 2, we used two playback experiment paradigms: interactive and static. The experiment consisted of three experimental conditions: (1) interactive, (2) static and (3) control.

The rationale of the interactive playback paradigm is to simulate a natural antiphonal vocal interaction by engaging subjects in reciprocal vocal exchanges. Similar interactive playback experiments have been successfully employed in anurans and songbirds (Dabelsteen and McGregor 1996; Schwartz 2001), but have not previously been attempted in a primate species. Our analyses from Experiment 1 suggested that on average phees produced within 5 s of the preceding phee are antiphonal responses, whereas phees produced 6–15 s following the preceding phee are spontaneous vocal productions. We used these intervals in designing this experiment. Because we have not experimentally tested the significance of latency to antiphonal call, we used the average timing intervals from Experiment 1 here. It is possible that our failure to adhere to context specific temporal patterns in antiphonal calling behavior may have decreased the number of antiphonal calls elicited by this playback procedure.

The logic behind our interactive playbacks was to try to mimic the subjects' natural vocal behavior from Experiment 1 by broadcasting 'antiphonal' phees from a hidden speaker each time the subjects produced a phee. However, if subjects did not produce a phee following a stimulus presentation, we broadcast a 'spontaneous' phee in hopes of eliciting an antiphonal response. The procedure for this paradigm was the following. We first waited for subjects to emit a phee, typically this occurred within the first 30 s the animal was in the test chamber. Once subjects emitted the initial phee, we broadcast a stimulus within 2–5 s of the offset of the subjects' vocalization (antiphonal stimulus presentation). Following the initial stimulus presentation each time the subjects produced a phee, we broadcast a stimulus within 2–5 s of the subjects' vocalization (antiphonal stimulus presentation). If subjects did not produce a phee within 15 s (±2 s) of an antiphonal stimulus presentation, we broadcast a stimulus (spontaneous stimulus presentation). This procedure continued until we broadcast 20 stimuli, at which point the experiment was ended.

The static paradigm contrasted with the 'interactive' playbacks in the following ways. Rather than broadcasting stimuli in response to the subjects' phees, we broadcast each stimulus at a set 15 s interval independent of the subjects' vocal behavior. As in the 'interactive' paradigm, the experiment continued until all 20 stimuli were broadcast.

We employed the interactive playback paradigm for both the 'interactive' and 'control' conditions, and the static playback paradigm in the 'static' condition. For the 'interactive' and 'static' conditions, we used phees produced by the subjects' CM, while in the 'control' condition the stimulus was a 'silent' sound file of the same duration as an average phee call stimulus (2.5 s). For both the 'interactive' and 'static' conditions, the stimulus set consisted of ten different exemplars of antiphonally produced phees from the subjects' CM recorded in Experiment 1. Each of these was presented twice in a random order. We broadcast the phees produced by the subjects' pair bonded mate in all cases except for one. This individual phees broadcast was of an adult CM of the opposite sex. We ran subjects four times on each condition in a randomized order.

In all conditions, we broadcast stimuli from a computer through an audio interface (eMagic a6/2 m), amplifier (Crown Model D-75A) and a free-field speaker (Cambridge Soundworks M80, frequency range: 40–22,000 Hz). We broadcast all phee stimuli at ~90 dB SPL measured at 1 m from the speaker and recorded subjects' vocal behavior during each trial using a directional microphone (Sennheiser ME-66). All sounds were recorded to an Apple G4 laptop via an audio interface (Digidesign mBox).

Analysis

Results from Experiment 1 suggested that the subjects typically emitted phees in response to hearing a CM's phee within 5 s of hearing the initial call. We classified these phees as antiphonal calls. The subjects also produced a second class of calls that did not appear to be a response to the call. These spontaneously produced calls had a longer latency (Fig. 2). For this experiment, we considered all phees produced within 5 s of broadcasting a stimulus of an antiphonal call and all phees produced from 6–15 s of a spontaneous call. For each test session, we tallied the number of antiphonal and spontaneous calls. We compared the subjects' vocal responses within and between conditions using repeated-measure ANOVA.

Results

Effects of social context (experiment 1)

We recorded a total of 1,450 phee vocalizations during the course of these experiments. Of those calls 1,183 were produced when subjects were paired with another monkey in the experimental setup (i.e., CM, NCM-OS, NCM-SS), whereas subjects produced the remaining calls in the ALONE condition. During the paired test conditions, we classified 583 phees as 'between-individual' and 600 as 'within-individual'. Although we recorded all vocal behavior that occurred during the test session, our analyses here focus on the patterns of vocal behavior that occurred within 30 s of each produced call. The reason for this is that 94% of all phees produced were within 30 s of the preceding phee. We considered calls produced after 30 s to be outliers.

Overall, results indicated that 'between-individual' (4.91 s, SE = 0.21) calls had a significantly shorter mean latency relative to 'within-individual' phees (10.23 s, SE = 0.26; $t(1104) = 15.84$, $P < 0.0001$; Fig. 2a). As social context could potentially affect aspects of subjects' vocal behavior, we directly compared both 'within-individual' and 'between-individual' calls in each of these contexts (Fig. 2b). Overall, the subjects showed a significantly shorter latency for 'between-individual' calls than for 'within-individual' calls (CM—'between-individual': mean = 4.6 s, SE = 0.30; 'within-individual': mean = 11.29 s, SE = 0.48 – $t(382) = 12.22$, $P < 0.0001$; NCM-SS—'between-individual': mean = 5.61 s, SE = 0.46; 'within-individual': mean = 9.60 s, SE = 0.38 – $t(384) = 6.69$, $P < 0.0001$; NCM-OS—'between-individual': mean = 4.71 s, SE = 0.37; 'within-individual': mean = 9.93 s, SE = 0.51 – $t(335) = 8.39$, $P < 0.0001$). Whereas subjects showed no statistical difference in latency to produce 'between-individual' calls between any of the conditions, 'within-individual' calls did vary between conditions. The subjects' latency to produce 'within-individual' calls clustered into two groups. The latency to call was similar in the CM (mean = 11.29s, SE = 0.48) and ALONE (mean = 11.76s, SE = 0.48) conditions, while the NCM-SS (mean = 9.60s, SE = 0.38) and NCM-OS (mean = 9.93s, SE = 0.51) conditions were also statistically indistinguishable. But all of the

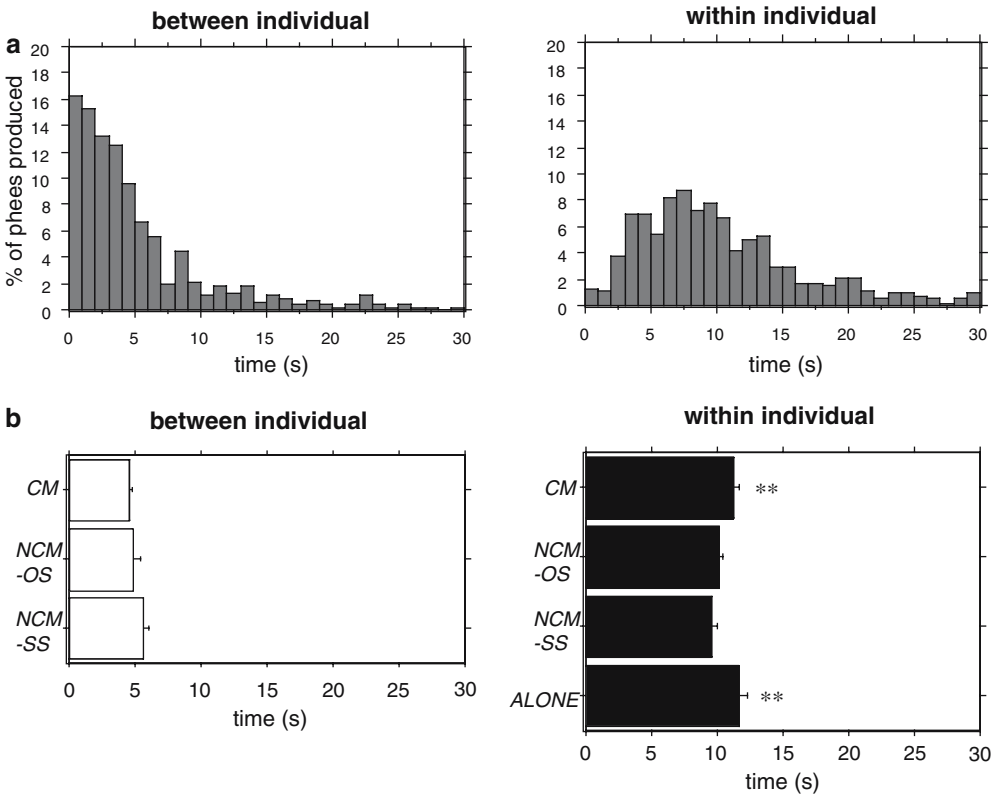

Fig. 2 Comparison between 'within-individual' and 'between-individual' phees. **a** Histograms show the proportion of calls produced within 30 s of the preceding vocalization. **b** *Bar* graphs show the mean (*standard error bars*) latency to emit 'between-individual' (left), and 'within-individual' phees for each of the social contexts subjects were recorded. *CM* cagemate, *NCM-OS* non-cagemate opposite sex, *NCM-SS* non-cagemate same sex; *ALONE* isolated calling (** denotes statistical difference)

comparisons between these conditions revealed statistical differences with CM and ALONE showing longer latencies (CM×NCM-SS: $t(380) = 2.79$, $P = 0.005$; CM×NCM-OS: $t(323) = 1.94$, $P = 0.05$; ALONE×NCM-SS: $t(406) = 3.43$, $P = 0.0007$; ALONE×NCM-OS: $t(349) = 2.48$, $P = 0.01$; Fig. 2b).

Analyses of the number of 'between-individual' calls produced in these experiments indicated no difference in the subjects' propensity to produce these calls in any of the experimental conditions (Fig. 3). Overall, the mean number of 'between-individual' calls produced in a test session was relatively consistent, with 10.19 produced during CM sessions, 8.48 during NCM-SS test sessions, and 8.81 during NCM-OS sessions. No statistical difference emerged between these conditions ($F(2, 5) = 0.12$, $P = 0.88$) and no interaction between condition and test session was evident ($F(4, 3) = 4.93$, $P = 0.11$) suggesting that the subjects' vocal behavior was consistent throughout the course of these experiments.

Previous studies reported sex differences in the acoustic structure of phees (Norcross and Newman 1993). To determine whether sex differences existed in marmoset vocal behavior during phee production, we conducted the following analyses. Collapsing all condi-

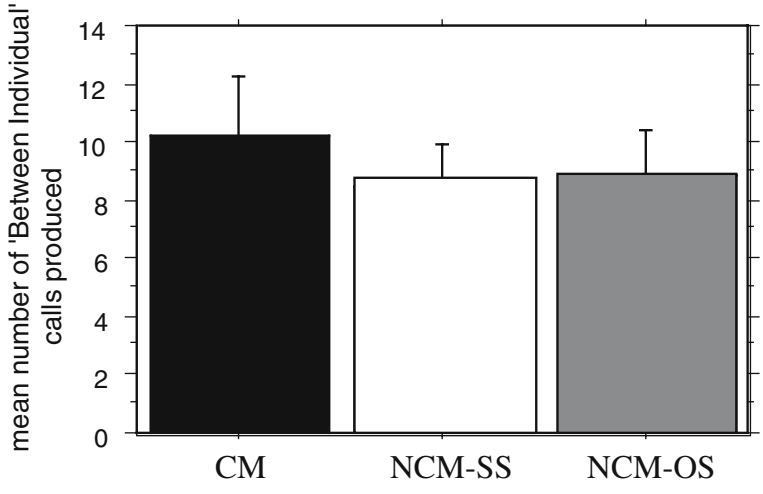

Fig. 3 Number of 'between-individual' calls produced in each of the social contexts. The *bar* graph depicts the mean number (*standard error bars*) of 'between-individual' calls produced in the cagemate (*CM*), non-cagemate same sex (*NCM-SS*) and non-cagemate opposite sex (*NCM-OS*) conditions

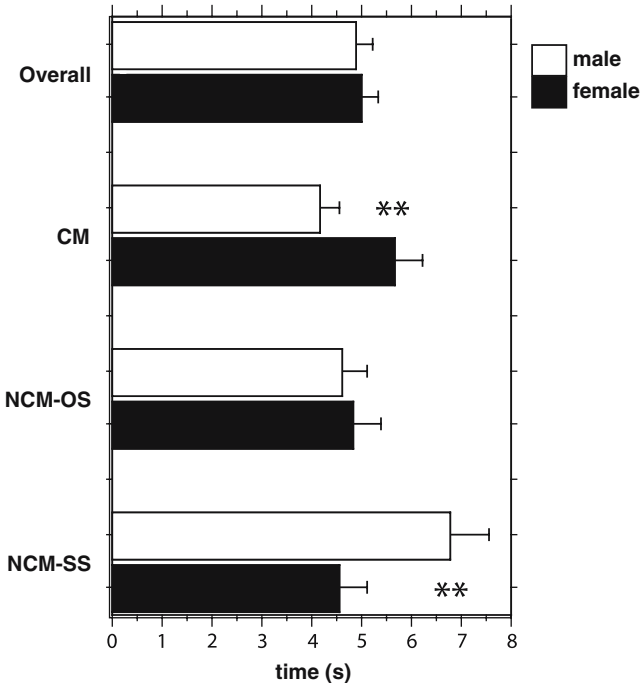

Fig. 4 Sex differences in marmoset vocal behavior. *Bar* graphs show the mean latency (*standard error bars*) to produce 'between-individual' phees for males (*white bars*) and females (*black bars*). Shown at the top is a summary of all conditions (Overall), as well as the results for each individual condition: cagemate (*CM*), non-cagemate same sex (*NCM-SS*) and non-cagemate opposite sex (*NCM-OS*). (** denotes statistical difference)

tions, no sex differences were evident in mean latency to produce 'between-individual' phees (male: mean = 4.91 s, SE = 0.29; female: mean = 4.99 s, SE = 0.32; Fig. 4). Differences did, however, emerge within specific conditions. Male subjects in the CM condition had a significantly shorter mean latency (4.18 s, SE = 0.35) than females (5.67 s, SE = 0.57; $t(226) = 2.37$, $P = 0.01$). In the NCM-SS condition, females had a significantly shorter mean latency (4.58 s, SE = 0.50) than males (6.78 s, SE = 0.78; $t(172) = 2.47$, $P = 0.01$). In contrast to these conditions, males and females showed no difference in vocal behavior during the NCM-OS condition (male: 4.6 s, SE = 0.49; female: 4.8 s, SE = 0.57). These results are shown in Fig. 4.

We observed that the subjects frequently produced 'between-individual' calls in a series, or bouts, of reciprocal calls. To test whether there were any differences in the length of these bouts between the different conditions, we calculated the number of calls that occurred in a consecutive sequence (Fig. 5). We defined these call sequences as a repetition of 'between-individual' calls, which by definition involved alternating phee productions by two monkeys (i.e., monkey A, monkey B, monkey A, monkey B, etc). This analysis is for the total number of consecutive calls produced by both monkeys engaged in the vocal interaction. Results indicated that mean length of 'between-individual' sequences was 3.7 calls in the CM condition, 2.5 calls in the NCM-SS condition, and 2.6 calls in the NCM-OS condition. Whereas the length was significantly longer in the CM condition than in either the NCM-SS ($t(127) = 2.17$, $P = 0.03$) or NCM-OS ($t(122) = 1.99$, $P = 0.04$) conditions, there was no difference between the two NCM conditions. A further difference between the conditions was the range in the length of the sequences. During the CM condition, the subjects had a range from 1–25 consecutive reciprocal calls. But in the NCM conditions, the length was much shorter, with 11 being the longest sequence in the NCM-SS condition and eight for the NCM-OS condition.

As a final analysis, we examined whether males or females were more likely to initiate bouts of reciprocal calls. Across all conditions, female subjects initiated an average of 34.3 bouts, while males initiated an average of 21.3 bouts. This differences, however, was not statistically significant due to significant individual variability ($F(2, 4) = 0.43$, $P = 0.68$). Similarly, when we analyzed individual conditions, no differences emerged. Although some individual differences were evident, there did not seem to be any consistent patterns.

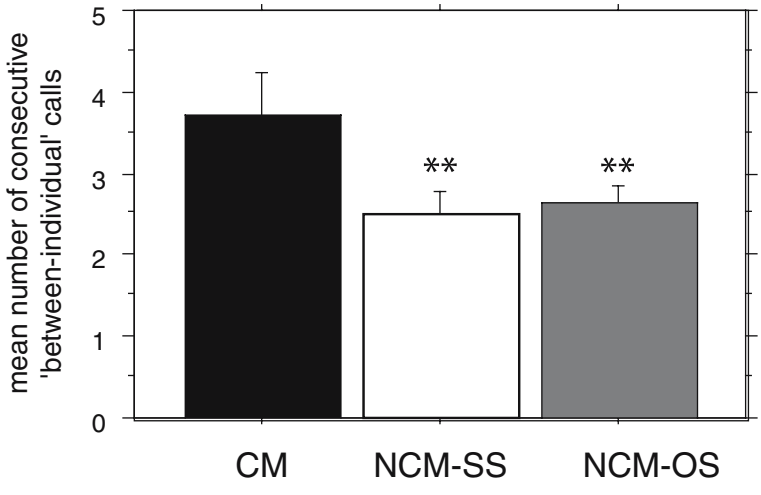

Fig. 5 Length of reciprocal calling sequences. The *bar* graph shows the mean number (*standard error bars*) of consecutive 'between-individual' calls produced in three experimental conditions: cagemate (*CM*), non-cagemate same sex (*NCM-SS*) and non-cagemate opposite sex (*NCM-OS*). (** denotes statistical difference)

Experimentally induced antiphonal calling (experiment 2)

Using the results from Experiment 1, we defined *antiphonal* calls as phees produced within 5 s of the stimulus and *spontaneous* calls as all phees produced 6–15 s following a stimulus. Results from our playback experiments indicated a significant difference in the number of antiphonal calls produced in the different experimental conditions (Fig. 6a). The subjects produced significantly more antiphonal calls in the interactive condition than in either the static ($F(1, 6) = 17.88$, $P = 0.006$) or control ($F(1, 6) = 15.65$, $P = 0.007$). There was, however, no difference between the static and control conditions ($F(1, 6) = 0.03$, $P = 0.86$). No interaction between the condition and test session emerged in any of these analyses suggesting no difference in the pattern of responses across the experiments.

Analyses of the subjects' spontaneous vocal behavior during this experiment also revealed several statistical differences (Fig. 6b). Although no difference emerged in the number of spontaneous calls produced in the interactive and control conditions ($F(1, 7) = 1.38$, $P = 0.28$), subjects did produce significantly fewer spontaneous calls in the static condition than in either the interactive ($F(1, 7) = 11.97$, $P = 0.01$) or control ($F(1, 7) = 1.38$, $P = 0.28$). Results revealed no interaction between condition and test session was evident in any of these comparisons.

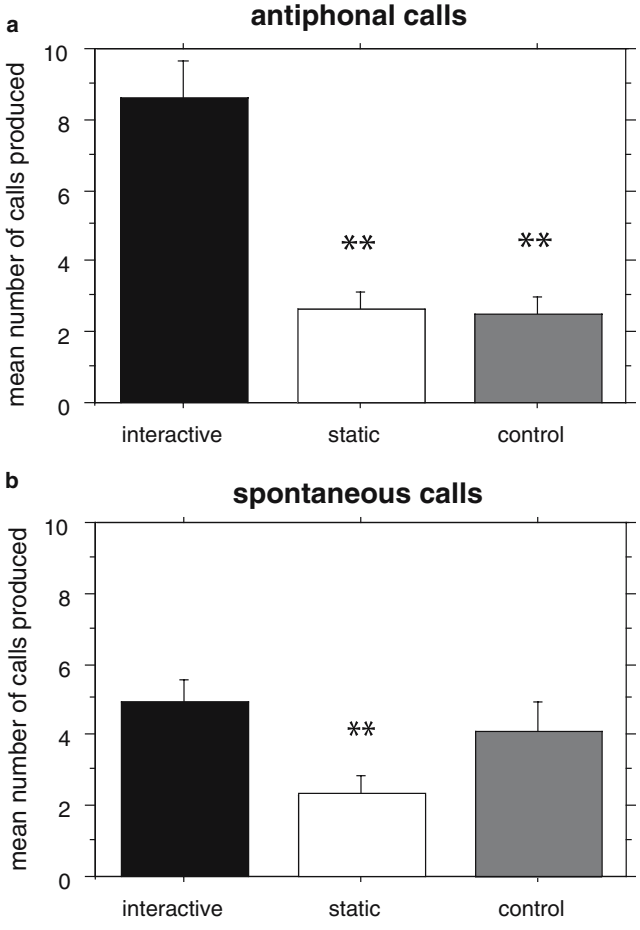

Fig. 6 Results from the playback experiments. **a** The *bar* graph here shows the mean number (*standard error bars*) of antiphonal calls produced in each of the three stimulus conditions. **b** The *bar* graph here shows the mean number (*standard error bars*) of spontaneous calls produced in each of the three stimulus conditions (** denotes statistical difference)

One important goal in performing the playback experiments was to determine whether we could design a paradigm that closely resembled natural antiphonal calling. To test this, we compared the vocal behavior exhibited in the 'interactive' playback condition and compared it to the CM condition from the first experiment across two dimensions: the number of consecutive calls and the total number of antiphonal or 'between-individual' calls produced. We chose these conditions for these analyses because the 'interactive' condition of the second experiment was designed to mimic the CM condition. Analyses revealed that the mean number of consecutive antiphonal calls in the 'interactive' playback experiment was 3.6, while the mean number of 'between-individual' calls in the CM condition was 3.7. This difference was not statistically different (Fig. 7a). The mean number of antiphonal calls produced in the 'interactive' playback condition was 8.6. In the CM condition, subjects produced a mean of 10.1 'between-individual' calls. Statistical analyses showed no difference between these two conditions (Fig. 7b).

Discussion

Here we investigated antiphonal calling in common marmosets to test how changes in the sensory input would affect the motor component of this natural vocal behavior. In Experiment 1, we recorded marmosets' natural vocal behavior and examined how changes in social context affected phee production. Experiment 2 built on this result in a playback experiment and tested explicitly how changes to the temporal interval between stimulus presentations affected antiphonal calling. We hypothesized that manipulation of the sensory input would modulate motor output in at least two ways: subjects' overall volubility and the timing of vocal production.

The first experiment sought to characterize the temporal relationship between the sensory and motor components of antiphonal calling. For this part of the study, we recorded the natural, long distance vocal behavior of marmosets in four contexts. Data showed that the timing of phee production clustered into two groups depending on the preceding vocal behavior. Those phees for which the preceding phee was produced by the other individual in the experiment, or ''between-individual' calls, had a significantly shorter latency, with 68% of calls being produced within 5 s of the previous phee and 78% within 7 s (Fig. 2a). In contrast, calls for

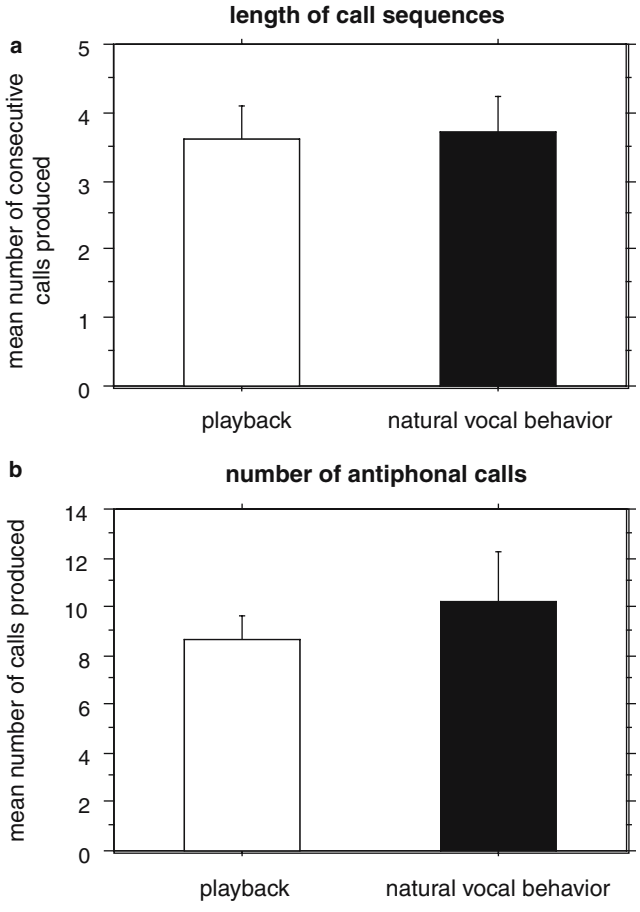

Fig. 7 Comparison between playback experiment and natural vocal behavior. **a** *Bar* graph shows the mean number (*standard error bars*) of consecutive antiphonal calls produced in the 'interactive' playback condition (playback) and 'between-individual' calls produced during the cagemate condition (natural vocal behavior). **b** *Bar* graph depicts the mean number (*standard error*) of antiphonal calls produced in the 'interactive' playback condition (playback) and 'between-individual' calls produced in the cagemate condition (natural vocal behavior)

which the preceding phee was produced by the same individual, or 'within-individual' calls, had a comparatively longer latency. Only 18% of 'within-individual' calls were produced within 5 s of the previous phee, but subjects produced 62% of these phees between 5–15 s (Fig. 2a). Based on these results, we developed the working definition that antiphonal calls phees produced within ~5 s of a conspecific's phee, whereas phees produced after this period of time are spontaneous calls likely emitted to induce an antiphonal response from a conspecific. However, results from our analyses of contextual variation in the timing of 'between-individual' calling (Fig. 2b) suggest that the timing of antiphonal calls is more dynamic.

The subjects' latency to produce 'between-individual' phees in the different test conditions was largely constant, but further analyses revealed significant sex differences in this aspect of vocal production. In the CM condition, for example, males had a significantly shorter latency to produce 'between-individual' phees than their female mate. However, when the male and female were from different cages (i.e., NCM-OS condition), no difference in vocal behavior occurred. During the NCM-SS condition, again sex differences emerged. But here males had a longer latency to call than females. The changes in call timing likely have communicative significance, with subjects inhibiting the vocal response for different periods of time depending on the sex and social relationship of the other individual. Exactly what is being communicated in these latency differences cannot be ascertained by the current dataset, but will be the subject of future investigations. This analysis suggests that what constitutes an antiphonal call may be context dependent, with some calls being antiphonal up to 7 s following the preceding phee. The significance of latency to antiphonal call in different social contexts will need to be examined in playback experiments to elucidate its communicative significance.

In addition to having an effect on the timing of phee production, the sensory input also induced modulations in other aspects of the subjects' vocal behavior. As shown in Fig. 1b, antiphonal calls frequently occur in sequences of reciprocal vocal exchanges. Although the number of 'between-individual' calls produced in the various conditions did not differ between conditions (Fig. 3), data indicated that the number of consecutive calls was variable. Specifically, the sequences in the CM condition consisted of significantly more consecutive calls than in either the NCM-SS or NCM-OS conditions (Fig. 5b). Since subjects are visually occluded throughout the experiment, these data suggest that subjects are able to recognize the identity of the caller and modify their vocal behavior accordingly. As such, the phee likely provides at least two types of information, the presence of a conspecific and the identity of that individual, each of which modulates different aspects of the motor output. Because the longest call sequences were between CMs, it suggests that the length of these vocal exchanges indicates the subjects' willingness to communicate.

Our analyses suggest that the subjects were able to recognize the other individual involved in the paired conditions of Experiment 1. Yet before we can conclude that the subjects' recognition was solely based on the vocalization, it is important to consider whether other sensory cues might contribute to this process. During transport from the colony and throughout the experiment subjects were visually occluded and therefore could not have used any visual cues to identify the other animal. One sensory modality, however, that could have been used was olfaction. Marmosets are known to use olfactory cues in a variety of social behaviors and these scents are likely to encode some information about individual identity (Epple et al. 1993; Smith et al. 2001). Although we cannot completely rule out the possibility that olfactory cues were used for individual recognition during these experiments, it seems unlikely for the following two reasons.

First, we cleaned all cages and transport boxes with a solution that is very effective at eliminating odors. This would presumably weaken any scents produced by subjects during the experiments. Indeed over time, the effectiveness of this solution does decrease. But subjects were only in the test room for 15 min during Experiment 1. Second, individual vocal recognition is common among primates (Cheney and Seyfarth 1982; Rendall et al. 1996; Bergman et al. 2003), including other Callitrichid species (Miller et al. 2001b; Weiss et al. 2001). As such, even if olfactory cues could be used to recognize the individual identity of another marmoset, it is likely that subjects could discern this information from the phees alone. In all, it may be that both cues could contribute to recognition, but given the experimental setup here we contend that recognition was based on acoustic cues in these experiments.

Building on the results of the first experiment, we employed playback experiments to test the significance of temporal intervals in the natural antiphonal calling observed in Experiment 1 as a sensory cue for eliciting this vocal behavior. For these experiments, two playback paradigms were used: interactive and static. During the interactive paradigm, stimulus presentation was dependent on the subjects' vocal behavior. The aim was to engage the subjects in natural reciprocal antiphonal vocal exchanges using the timing intervals recorded in the first experiment. For the static paradigm, in contrast, we presented stimuli at a constant temporal interval (15 s) that was independent of the subjects' vocal behavior. Critically, the total number of stimuli we presented was identical in all conditions. If reciprocity is communicatively significant for antiphonal calling, we should observe a significant decrease in the subjects' vocal behavior between the 'interactive' and 'static' conditions. If, however, reciprocity does not serve as a communication cue in antiphonal calling, then no differences will emerge between the two conditions.

Subjects produced significantly more antiphonal calls in the 'interactive' playback condition than in the 'static' condition (Fig. 6a). In fact, subjects produced the same number of antiphonal calls during the static condition as when we presented subjects with 'silent' stimuli in the control condition. Further analyses showed that in addition to eliciting fewer antiphonal calls, subjects produced significantly fewer spontaneous calls in the static condition relative to both the interactive and control conditions. Hence subjects were actively inhibiting their overall vocal production as a result of the static stimulus presentation. This suggests that reciprocity in antiphonal vocal exchanges is an important cue that indicates an individual's intention to communicate with a particular conspecific.

Despite being a stereotyped vocal behavior, the data presented here indicate that antiphonal calling in the common marmoset is, in fact, quite dynamic. Changes to several aspects of the sensory input, such as the presence of a conspecific, the identity and sex of the caller, as well as whether the occluded individual participates in reciprocal antiphonal exchanges, result in modulations of the vocal output. The observed changes in the motor output, however, are not binary (i.e., call–no call). Vocal production varied in the overall number of calls produced, including inhibition of vocal responses, the timing of production, and length of the antiphonal vocal exchanges. These data suggest that antiphonal calling in the common marmosets is sufficiently structured, so as to permit its use as an experimental assay, but can also be modulated along several predictable dimensions that are amenable to neurobiological inquiry.

Over the past several decades, a series of studies investigated the response properties of neurons in auditory cortex to species-specific primate vocalizations (Newman and Wollberg 1973; Winter and Funkenstein 1973; Rauschecker et al. 1995; Wang et al. 1995). More recently, however, attempts to characterize the response properties of prefrontal cortex neurons to primate vocalizations have taken place (Romanski and Goldman-Rakic 2001; Romanski et al. 2004). Interest in the role of frontal cortex in audition has occurred because this region of cortex is implicated in an array of high-level perceptual and cognitive processes from studies of other sensory modalities (Miller et al. 2003; Romo and Salinas 2003). Following work in the visual and somatosensory systems (Romo et al. 1999, 2004; Freedman et al. 2001; Nieder et al. 2002), a sensory-motor task is likely needed to probe the analogous mechanisms in audition. In these tasks, subjects are presented with a sensory stimulus and must retain that information for a period of time until they make a decision about how to respond using a motor behavior. As a sensory-motor behavior, antiphonal calling involves this exact sequence of events.

Antiphonal calling represents an ideal behavior to explore the role of neurons in both the auditory and frontal cortices during sensory-motor interactions. The properties of neurons in primary auditory cortex in response to both hearing (Wang et al. 1995; Wang 2000; Wang and Kadia 2001) and producing (Eliades and Wang 2003, 2005) vocalizations have been explored in the common marmoset. But to advance this line of research to vocal signal processing in the auditory belt and frontal cortex, it is critical to incorporate the use of behavior. Results presented here show that many of the cognitive functions implicated in the frontal cortex from studies of the visual system are necessary for antiphonal calling. Namely, marmosets must identify and categorize the identity of the conspecific caller and make a decision about whether to produce an antiphonal call as well as the timing of that vocalization. Recent technological advances make it possible to record neural activity in freely moving and vocalizing animals (Yu and Margoliash 1996; Dave et al. 1998; Fee and Leonardo 2001; Hahnloser et al. 2002; Luo et al. 2003; Leonardo and Fee 2005). Combining such technology with the antiphonal calling behavior will yield an exciting line of

research capable of addressing many of the questions about sensory-motor interactions in vocal communication, far beyond what is already known about acoustic signal processing in the marmoset auditory cortex specifically and primates as a whole.

Acknowledgements We thank Audrey DiMauro, Kate Mandel, Kate Rydstrom, Ambica Tumkur, and Ray Wang for their help in running these experiments. This research was supported by NIH grants F32 DC007022 (CTM) and R01 DC 005808 (XW). These experiments comply with the 'Principles of Animal Care', Publication No. 86–23, revised 1985, of the National Institutes of Health. All experimental protocols were approved by the Johns Hopkins University Animal Use and Care Committee.

References

Barbour D, Wang X (2003a) Auditory cortical responses elicited in awake primates by random spectrum stimuli. J Neuro 23:7194–7206

Barbour D, Wang X (2003b) Contrast tuning in auditory cortex. Science 299

Bergman TJ, Beehner JC, Cheney DL, Seyfarth RM (2003) Hierarchical classification by rank and kinship in baboons. Science 302:1234–1236

Carew TJ (2000) Behavioral neurobiology: the cellular organization of natural behavior. Sinauer Associates, New York

Cheney DL, Seyfarth RM (1982) Recognition of individuals within and between groups of free ranging vervet monkeys. Am Zool 22:519–529

Dabelsteen T, McGregor PK (1996) Dynamic acoustic communication and interactive playback. In: Kroodsma DE, Miller EH (eds), Ecology and evolution of acoustic commmunication in birds. Cornell University Press, Ithaca, pp 398–408

Dave A, Yu AC, Margoliash D (1998) Behavioral state modulation of auditory activity in a vocal motor system. Science 282:2250–2254

Eliades SJ, Wang X (2003) Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J Neurophysiol 89:2185–2207

Eliades SJ, Wang X (2005) Dynamics of auditory-vocal interaction in monkey auditory cortex. Cerebral Cortex (In Press)

Epple G, Belcher AM, Kuderling I, Zeller U, Scolnock L, Greenfield KL, Smith III AB (1993) Making sense out of scents: species differences in scent glands, scent-marking behavior and scent mark composition in the Callitrichidae. In: Rylands AB (ed) Marmosets and tamarins: systematics, behaviour and ecology. Oxford University Press, Oxford, pp 123–153

Fee MS, Leonardo A (2001) Miniature motorized microdrive and commutator system for chronic neural recording in a small animal. J Neuro Methods 112:83–94

Freedman DJ, Anderson KC, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. Science 291:312–316

Ghazanfar AA, Hauser MD (2001) The auditory behaviour of primates: a neuroethological perspective. Curr Op Neuro 11:712–720

Ghazanfar AA, Flombaum JI, Miller CT, Hauser MD (2001) The units of peception in cotton-top tamarin (*Saguinus oedipus*) long calls. J Comp Physiol A 187:27–35

Ghazanfar AA, Smith-Rohrberg D, Pollen A, Hauser MD (2002) Temporal cues in the antiphonal calling behaviour of cotton-top tamarins. Anim Behav 64:427–438

Grace JA, Amin N, Singh NC, Theunissen FE (2003) Selectivity for conspecific song in the zebra finch auditory forebrain. J Neurophysiol 89:472–487

Hahnloser RHR, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419:65–70

Hauser MD (1996) The evolution of communication. MIT Press, Cambridge

Hessler N, Doupe AJ (2000) Social context modulates singing-related neural activity in the songbird forebrain. Nat Neuro 2:209–211

Jarvis ED, Scharff C, Grossman M, Ramos JA, Nottebohm F (1998) For whom the bird sings: context-dependent gene expression. Neuron 21:775–788

Leonardo A, Fee MS (2005) Ensemble coding of vocal control in birdsong. J Neuro 25:652–661

Liberman AM (1996) Speech: a special code. MIT press, Cambridge

Lu T, Liang L, Wang X (2001a) Temporal and rate representations in time varying signals in the auditory cortex of awake primates. Nat Neuro 4:1131–1138

Lu T, Liang L, Wang X (2001b) Neural representation of temporally asymmetric stimuli in the auditory cortex of awake primates. J Neurophysiol 85:2364–2380

Luo M, Fee MS, Katz LC (2003) Encoding phermonal signals in the accessory olfactory bulb of behaving mice. Science 299:1196–1201

Marler P, Slabbekoorn H (2004) Nature's music: the science of birdsong. Elsevier/Academic Press, Amsterdam

Miller CT, Hauser MD (2004) Multiple acoustic features underlie vocal signal recognition in tamarins: antiphonal calling experiments. J Comp Physiol A 190:7–19

Miller CT, Dibble E, Hauser MD (2001a) Amodal completion of acoustic signals by a nonhuman primate. Nat Neuro 4:783–784

Miller CT, Miller J, Costa RGD, Hauser MD (2001b) Selective phontaxis by cotton-top tamarins (*Saguinus oeidpus*). Behaviour 138:811–826

Miller CT, Scarl JS, Hauser MD (2004) Sensory biases underlie sex differences in tamarin long call structure. Anim Beh 68:713–720

Miller CT, Iguina C, Hauser MD (2005) Processing vocal signal for recognition during antiphonal calling. Anim Beh 69:1387–1398

Miller EK, Nieder A, Freedman DJ, Wallis JD (2003) Neural correlates of categories and concepts. Curr Op Neurobiol 13:198–203

Newman JD, Wollberg Z (1973) Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. Brain Res 54:287–304

Nieder A, Freedman DJ, Miller EK (2002) Representation of quantity of visual items in the primate prefrontal cortex. Science 297:1708–1711

Norcross JL, Newman JD (1993) Context and gender specific differences in the acoustic structure of common marmoset (*Callithrix jacchus*) phee calls. Am J Prim 30:37–54

Norcross JL, Newman JD (1997) Social context affects phee call production by nonreproductive common marmosets (*Callithrix jacchus*). Am J Prim 43:135–146

Norcross JL, Newman JD, Fitch WT (1994) Responses to natural and synthetic phee calls by common marmosets. Am J Prim 33:15–29

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, *Serinus canarius*. J Comp Neurol 165:457–468

Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268:111–114

Rendall D, Rodman PS, Edmond RE (1996) Vocal recognition of individuals and kin in free-ranging rhesus monkeys. Anim Beh 51:1007–1015

Romanski LM, Goldman-Rakic PS (2001) An auditory domain in primate prefrontal. Nat Neuro 5:15–16

Romanski LM, Averbeck BB, Diltz M (2004) Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. J Neurophysiol 93:734–747

Romo R, Salinas E (2003) Flutter discrimination: neural codes, perception and decision making. Nat Rev Neuro 4:203–221

Romo R, Brody CD, Hernandez A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. Nature 399:470–473

Romo R, Hernandez A, Zainos A (2004) Neuronal correlates of a perceptual decision in ventral premotor. Neuron 41:165–173
Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. Nat Neuro 1:513–518
Schwartz JH (2001) Call monitoring and interactive playback systems in the study of acoustic interactions among male anurans. In: Ryan MJ (eds) Anuran communication. Smithsonian Institution Press, Washington
Sen K, Theunissen FE, Doupe AJ (2001) Feature analysis of natural sounds in the songbird auditory cortex. J Neurophysiol 86:1145–1458
Smith TE, Tomlinson AJ, Mlotkiewcz JZ, Abbott DH (2001) Female marmoset monkeys (*Callithrix jacchus*) can be identified from chemical composition of their scent marks. Chem Sens 26:449–458
Spiro JE, Dalva MB, Mooney R (1999) Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. J Neurophysiol 81:3007–3020
Wang X (2000) On cortical coding of vocal communication sounds in primates. Proc Nat Acad Sci 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) Representations of species-specific vocalizations in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol 74:2685–2706
Weiss DJ, Garibaldi BT, Hauser MD (2001) The production and perception of long calls by cotton-top tamarins (*Saguinus oedipus)*: acoustic analyses and playback experiments. J Comp Psych 11:258–271
Winter P, Funkenstein HH (1973) The effect of species-specific vocalizations on the discharge of auditory cortical cells in the awake squirrel monkey. Exp Brain Res 18:489–504
Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. Science 273:1871–1874