J Comp Physiol A (2004) 190: 7–19
DOI 10.1007/s00359-003-0468-1

ORIGINAL PAPER

C. T. Miller · M. D. Hauser

Multiple acoustic features underlie vocal signal recognition in tamarins: antiphonal calling experiments

Received: 18 February 2003 / Revised: 11 September 2003 / Accepted: 16 October 2003 / Published online: 11 November 2003

Abstract We examine which acoustic features are relevant for recognition of the cotton-top tamarin (*Saguinus oedipus*) combination long-call. This vocalization, emitted by both males and females, functions in maintaining group cohesion, territory defense and mate attraction. Using the tamarins' natural antiphonal vocal response to hearing a combination long-call as the primary measure of recognition, we presented subjects with synthetic exemplars of combination long-calls in which we manipulated across one of three acoustic dimensions: frequency, time and amplitude. Results indicated that although acoustic features in the frequency and time domains are important for combination long-call recognition, the changes in amplitude within and between syllables are not. Furthermore, while the fundamental frequency appears to be the used to encode information about the frequency contour, the temporal information is derived from the harmonics. Overall, these results suggest that tamarins use a specific suite of acoustic features for combination long-call recognition.

Keywords Antiphonal calling · Cotton-top tamarins · Vocal signal recognition

Abbreviations *CLC* combination long-call · *F0* fundamental frequency · *JND* just noticeable difference · *JMD* just meaningful difference · *IPI* inter-pulse interval

C. T. Miller · M. D. Hauser
Primate Cognitive Neuroscience Laboratory,
Department of Psychology, Harvard University

Present address: C. T. Miller
Laboratory of Auditory Neurophysiology,
Department of Biomedical Engineering,
Johns Hopkins University School of Medicine,
420 Rutland Ave., Baltimore, MD, 21205 USA
E-mail: ctmiller@bme.jhu.edu
Tel.: +1-410-5026968
Fax: +1-410-6149599

Introduction

All species that mediate behavioral interactions between conspecifics with vocal signals must possess auditory systems that can distinguish between sounds produced by conspecifics and other sounds in the environment. Studies of songbirds and anurans indicate that auditory recognition systems are tuned to detect a specific combination of acoustic characteristics unique to conspecific vocalizations (Nelson 1988; Gerhardt 1991; Searcy et al. 1995). At the neural level, combination-sensitive neurons fire when such identifying acoustic features are detected, thus signaling the central nervous system that the sound is a conspecific signal (Doupe 1997; Whaling et al. 1997; Grace et al. 2003). While our understanding of this process is advanced in anurans (Ryan 2001; Gerhardt and Huber 2002) and songbirds (Doupe and Solis 1999), considerably less is known about vocal signal recognition at either the behavioral or neural level in nonhuman primates (hereafter primates). In the present study, we sought to reduce this gap by conducting a study of vocal signal recognition in a New World primate species, the cotton-top tamarin (*Saguinus oedipus*), using an ethologically relevant experimental assay: antiphonal calling (Ghazanfar et al. 2001, 2002; Miller et al. 2001).

The study of the neural basis of anuran and songbird vocal behavior is based on extensive research detailing the natural behavior of these species. In the case of song recognition, behavioral work has focused on song learning (Marler 1970, 1997; Nelson 1997; Soha and Marler 2000) and species recognition (Emlen 1972; Nelson 1988; Searcy and Brenowitz 1988). Building on this work, researchers have examined the neural mechanisms that underlie these aspects of song recognition (Doupe 1997; Solis and Doupe 1997, 2000; Whaling et al. 1997; Doupe and Solis 1999; Sen et al. 2001). Comparably little is known about the mechanisms that underlie vocal signal recognition in primates. However, this taxonomic group is particularly interesting because,

unlike songbirds and anurans, the neural representations of primate vocalizations are located in the neocortex (Newman and Wollberg 1973a, 1973b; Rauschecker et al. 1995; Wang and Kadia 2001). Given the unique anatomical organization of the neocortex, it is likely that these neural substrates employ different computations for recognition than in songbirds or anurans. Following the neuroethological approach taken in studies of songbird and anuran vocal behavior, employing ethologically relevant procedures to investigate vocal signal recognition in primates at the behavioral level will allow us to compare the neural mechanisms that underlie auditory recognition systems across taxonomic groups.

Cotton-top tamarins are highly vocal, small-bodied, New World monkeys that inhabit dense tropical rainforests. When visually occluded from group members, tamarins emit a long distance, species-specific vocalization known as the combination long call (CLC) (Miller et al. 2003a). The CLC is comprised of a concatenation of two acoustically distinct syllable types: one to two chirps followed by two to four whistles (Cleveland and Snowdon 1982; Fig. 1). Several studies have already contributed to an understanding of this call's acoustic morphology (Weiss et al. 2001), the mechanisms of vocal production (Miller et al. 2003b) and the factors influencing perceptual classification (Miller et al. 2003a). Upon hearing a CLC, tamarins typically respond by producing an acoustically similar CLC, a behavior referred to as antiphonal calling. Tamarin antiphonal calling has been the basis for several previous investigations of vocal signal recognition in tamarins (Miller et al. 2001; Ghazanfar et al. 2002). The logic of studies employing the antiphonal calling assay is straightforward. If subjects produce similar rates of antiphonal calls to playbacks of both a normal CLC and a manipulated CLC, then one can conclude that the manipulated feature, or features, was not important for call recognition in this context. If, however, subjects show a significant decrease in antiphonal calling to a manipulated stimulus, then one can conclude that the manipulated feature was important for some aspect of CLC recognition. Following studies of birdsong (Konishi 1985, 1994; Marler 1991), the motivation behind our studies of antiphonal calling is to develop a neuroethological experimental assay that can be used to probe the recognition system at both the behavioral and neural levels. These behavioral experiments will ultimately be essential to future studies examining the neural mechanisms underlying antiphonal calling.

Building on our earlier work, here we sought to investigate the role of multiple acoustic features in CLC recognition. Previous studies suggest that the antiphonal calling response in cotton-top tamarins is mediated by a combination of features across the structure of the entire CLC, rather than either syllable type alone (Ghazanfar et al. 2001). Further, aspects of the temporal pattern of the call, such as the amplitude envelope and inter-pulse interval, appear important for signal recognition (Ghazanfar et al. 2002). While these studies yield important insights into CLC recognition, there are still many questions left unanswered. Data suggest that the temporal pattern of CLCs is relevant for signal recognition, however it is not clear whether this information is encoded from the fundamental frequency (F0) or the harmonics. Further, no previous study has investigated the importance of the frequency contour or amplitude modulation within and between syllables for CLC recognition. The present study furthers our understanding of the tamarin auditory recognition system by systematically investigating the role of frequency, time and amplitude for CLC recognition.

To accomplish the goal of this investigation, it was important to use synthetic calls that could be manipulated along multiple acoustic dimensions. The use of synthetic vocalizations is widespread in studies of anurans and songbirds (Nelson 1988; Nelson and Marler 1990; Ryan 2001; Gerhart and Huber 2002), but only recently has been used in studies of nonhuman primates (Moody and Stebbins 1989; Owren 1990; Weiss and Hauser 2002). In order to use synthetic calls as test stimuli, it is necessary to show that the synthetic versions of calls are sufficient to elicit the desired behavior. Experiment 1 (natural versus synthetic) addresses this by comparing the antiphonal call responses of naturally produced CLCs with the synthetic replicas of these calls. Building on this result, in experiment 2 (signal

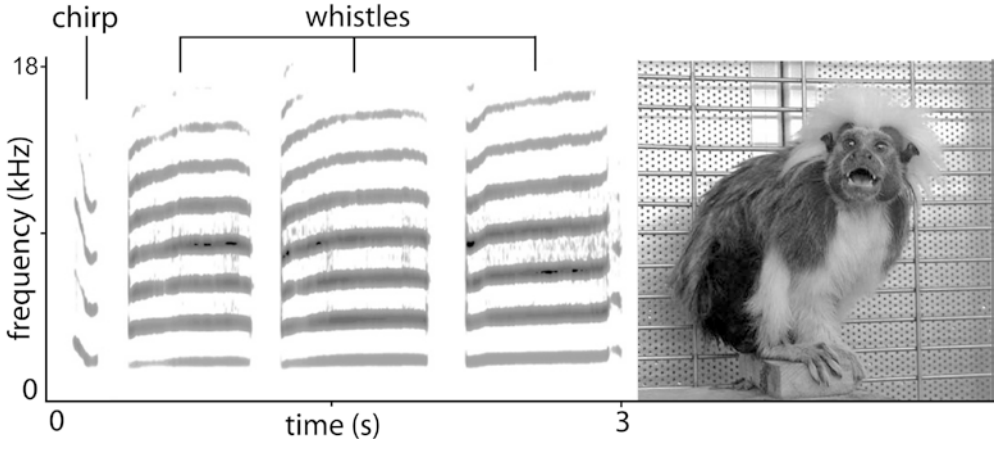

Fig. 1 A spectrogram of a 'combination long call' (CLC) is depicted to the *left* with syllable types denoted. To the *right* is an image of a cotton-top tamarin during production of this vocalization

manipulations) we examine the significance of particular acoustic features for CLC recognition across three acoustic dimensions: frequency, time and amplitude.

Materials and methods

Subjects

Subjects were adult cotton-top tamarins (*S. oedipus*) housed at the Harvard University Primate Cognitive Neuroscience Laboratory. All subjects were born in captivity, live in social groups consisting of a breeding pair, and in some cases one to two generations of offspring. Groups were provided with monkey chow, fruit, sunflower seeds, peanuts, and yogurt at the end of the day and have ad libitum access to water; during some experiments, they were given food rewards in the form of raisins or sweet cereal.

In experiment 1 (natural versus synthetic calls) we tested 13 subjects (6 male, 7 female). However, following the first test condition (normal call versus F0), we dropped 5 subjects from the experiments due to pregnancy ($n = 3$) or consistently failing to reach calling criteria ($n = 2$). We dropped one additional subject following the second test condition (natural F0 versus synthetic F0) due to illness. In experiment 2 (signal manipulations), 7 subjects (4 male, 3 female) participated in all test conditions.

Apparatus

The test apparatus consisted of a box with five Plexiglas sides and a wire mesh front placed inside an Industrial Acoustics Company (Model 400A) acoustic chamber. We broadcast all stimuli from an Apple G4 computer through either a Digidesign Audiomedia III soundcard or a Digidesign M-Box, both at 16-bit sound. Sound was broadcast through an Alesis RA-100 Amplifier and an Alesis Monitor One speaker (Frequency Range: 45–18,000 Hz). We monitored subjects' behavior using a video camera positioned inside the acoustic chamber and a video monitor outside the chamber. This system has been used in several previous studies of tamarin auditory perception (Ghazanfar et al. 2001, 2002; Miller et al. 2001).

Procedure

We removed subjects from the homeroom and brought them individually to the testing room via a transport box. Once inside the test chamber, subjects were placed inside the testing apparatus. We monitored subjects' activity outside the acoustic chamber from a camera located inside. Once inside the apparatus, the chamber was closed and the experiment started. All stimuli in the set were randomized and broadcast using a custom Hypercard program (written by W.T. Fitch) at ~30-s intervals. We broadcast all stimuli at approximately 65–70 dB SPL (measured 1 m from speaker location). We considered all instances in which subjects emitted a long call within 5 s of stimulus offset as an antiphonal call. This procedure follows previous studies of antiphonal calling (Ghazanfar et al. 2001, 2002). Following Ghazanfar et al. (2002), subjects were presented with the two stimulus sets in an alternating manner (i.e., set 1, set 2, set 1, set 2). Subjects were only run at most once each day. All testing occurred between 8:00 a.m. and 2:00 p.m. over a 16-month period.

Stimulus sets

For each test condition, we generated two stimulus sets. We presented each stimulus set to subjects twice for a total of four test sessions per condition. The organization of the stimulus sets for experiment 1 (natural versus synthetic calls) and experiment 2 (signal manipulations) differed. The reason for this was that experiment 1 was used to determine whether synthetic CLCs would elicit statistically similar rates of antiphonal calling. This was necessary for testing whether it was appropriate to use synthetic CLCs in the second experiment. Each stimulus set in experiment 1 consisted of 16 stimuli: 8 'normal' calls and 8 test stimuli. For experiment 2, we manipulated the acoustic structure of the signal along one of three dimensions: frequency, time, or amplitude. As there were multiple comparisons within each of these dimensions, it was important to present more than one test stimulus in each set. Stimulus sets in experiment 2 consisted of 20 total stimuli. Specifically, stimulus sets consisted of 5 exemplars of 'normal' CLCs and the three types of signal manipulations for each 'normal' CLC.

Stimulus generation

We generated synthetic stimuli for all conditions using a customized macro written by D. Smith-Rothberg and C. Miller for SIGNAL 4.0 (Beeman 2001). Similarly, we generated stimuli used in both frequency modulation conditions with macros written by C. Miller and R. Egnor for SIGNAL 4.0. Listed below is a description of how stimuli were created for each condition. We manipulated stimuli in the amplitude and time conditions using SoundEdit 16 (version 2) and CANARY (version 1.2).

All synthetic calls used in this experiment were generated from a natural CLC. Generating a synthetic signal from a natural call was done in the following way. Using SIGNAL, a spectrogram was generated and the frequency contour traced using the cursor. The F0 and any harmonics were traced and stored in separate buffers. The energy that occurs along the traced contour was stored for each harmonic. All of these buffers were then combined to create a final synthetic replica of the original signal.

In experiment 1 (condition 1) we created the F0-only stimulus by filtering out all harmonics in the signal. We created stimulus sets in experiment 1 by generating synthetic replicas of naturally produced calls. For experiment 1 (condition 2) the synthetic call was a replica of only the F0, while the synthetic stimulus used in experiment 1 (condition 3) consisted of a F0 as well as the first three harmonics. To determine that the synthetic calls used in experiment 1 were adequately similar the natural calls, we calculated spectrogram cross-correlations. For condition 2, the mean cross-correlation was 0.97 (SE = 0.009), while for condition 3 the mean cross-correlation was 0.94 (SE = 0.02). All stimuli used in experiment 2 were synthetic calls. The 'normal' call for each of these conditions was a synthetic replica of a natural call. We then manipulated synthetic calls for each of the 'normal' synthetic calls.

Experiment 1: synthetic versus natural calls

In this experiment, each stimulus set consisted of a total of 16 different stimuli. We selected two naturally produced CLC exemplars from 8 individuals. For each of the naturally produced calls, we generated a manipulated version of the call. The naturally produced call was always in the opposite stimulus set as the manipulated version of the call. In other words, the natural exemplar of CLC A might be in stimulus set 1, but the synthetic replica of that call would be in stimulus set 2.

Experiment 2: signal manipulations

All stimulus sets in experiment 2 consisted of 20 stimuli. We selected two naturally produced CLCs for 5 different individuals. A synthetic replica of each of these calls was generated and served as the baseline call for these conditions. Each synthetic replica was then manipulated in three different ways. The exact nature of the manipulations varied between test conditions (see below). All stimulus sets used in experiment 2 consisted of 5 normal calls produced by different individuals and 5 exemplars of three manipulations of those calls for a total of 20 stimuli. This stimulus design permitted us to compare an unmanipulated CLC across

several manipulated calls within a single condition. As in experiment 1, we arranged each stimulus set so that the 'normal' CLCs (i.e., unmanipulated synthetic calls) were in the opposite stimulus set from the manipulated versions of the call. We organized stimulus sets in this way to follow previous antiphonal calling experiments in our lab (Ghazanfar et al. 2001, 2002).

Stimulus manipulations

Frequency

Justification The frequency contour of tamarin CLCs shows modulation both within and between syllables. In contrast to many mammalian vocalizations (Hauser and Fowler 1992), the CLC, as well as other callitrichid long calls (Snowdon 1993), increase in frequency over the duration of the call (Weiss et al. 2001). Within each syllable the frequency typically increases slightly before tapering at the end of each syllable (Weiss et al. 2001). The following two sets of stimuli were used to test whether the modulation in the frequency contour is relevant for CLC recognition. These stimuli consisted of only a F0 contour.

Between-syllable frequency modulation In this condition, we manipulated the frequency modulation of test stimuli by generating stimuli that differed in the direction and amount of modulation between syllables, but with modulation within a syllable eliminated. Stimuli in this condition consisted of synthetic CLCs with only a F0. We generated the between-modulation stimulus by first measuring the start, peak and end frequency of each syllable in each unmanipulated CLC. The mean of these three measures was calculated for each chirp and whistle syllable. We then generated a CLC in which each syllable had the mean frequency of the syllable in the original signal. The no-modulation stimulus was created by generating whistles with unmodulated frequency contours positioned at the mean pitch (frequency) for the whole CLC. Stimuli in the reverse between-modulation condition were generated in the following manner. First, we measured the difference in frequency from the second whistle syllable to all other chirp and whistle syllables. Second, the mean frequency of the second whistle was manipulated to be the same as for the between-modulation stimulus. Third, each other chirp and whistle was manipulated to be the same frequency difference from the second whistle as in the between-modulation stimulus, but in the opposite direction. In other words, if whistle 1 was 300 Hz lower than whistle 2 in the between-modulation stimulus, then it would be 300 Hz higher in the reverse-modulation stimulus (see Fig. 2).

Within-syllable frequency modulation Here we manipulated the frequency contour within each syllable, while keeping the between-syllable modulation constant between all syllables. Stimuli in this condition consisted of synthetic CLCs with only a F0. The start, peak, and end frequency of each syllable was measured for each unmanipulated CLC. From these measurements, a mean frequency for the call was calculated. We created the within-modulation

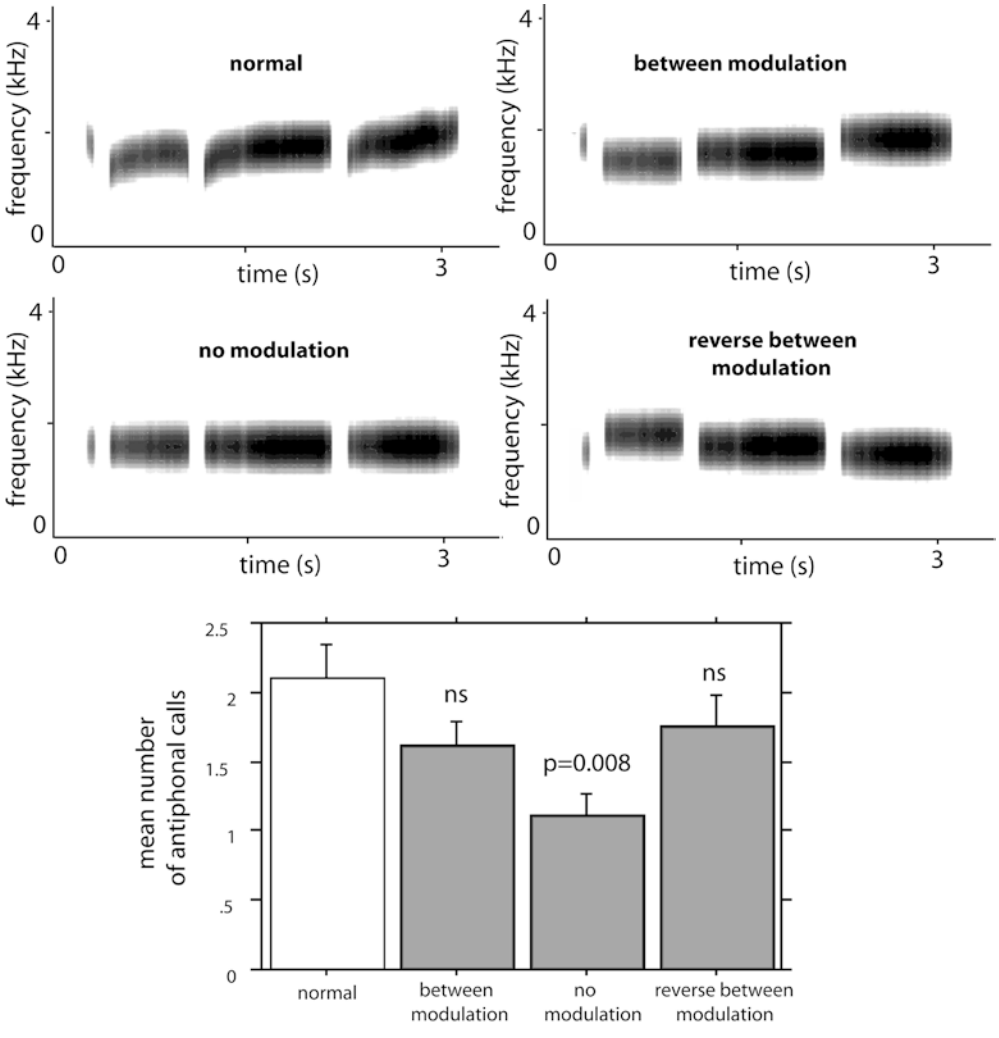

Fig. 2 Spectrograms of the four synthetic CLCs used in the between-frequency modulation test condition (experiment 2: condition 1) are shown at the *top*. *Below* is a bar graph showing the mean antiphonal calling responses for all stimuli with standard error bars. Antiphonal call responses to normal calls are shown *unshaded*, while responses to test stimuli are shown in *black*. For all stimulus types that elicited a statistically different antiphonal call rate from 'normal' calls, the *P* value is shown. (*ns*: not significantly different)

stimuli by generating whistles with the same frequency contour as in the unmanipulated CLC, but the mean frequency of each syllable was the same as the mean frequency of the whole call. The no-modulation stimulus was generated by creating whistles with unmodulated frequency contours positioned at the mean pitch for the whole CLC. We generated the *reverse-within modulation* stimuli by generating whistles with the inverse contour of whistles in the within-modulation stimuli (see Fig. 3).

Time

Justification The acoustic structure of CLCs consists of a series of pulses separated by inter-pulse intervals (IPIs). IPIs are known to be stable both within and between individuals (Miller et al. 2003b; Weiss et al. 2001) suggesting that they may be an important cue for signal recognition. Here we presented subjects with CLCs in which we manipulated the IPIs while leaving all other acoustic information intact.

Inter-pulse interval The stimuli for both conditions below were created by first generating synthetic CLCs consisting of only the F0 (condition 6: Fig. 4) or the F0 plus three harmonics (condition 7: Fig. 4). We measured the IPI between each syllable in the test CLCs. The IPI was measured as the duration of time (milliseconds) from the offset of one syllable to the onset of the subsequent syllable. For the *half-IPI* stimuli, we multiplied the IPI between each set of syllables by 0.5 and decreased each IPI to this duration. For *zero-IPI* stimuli, the IPI was eliminated. We created the *twice-IPI* stimuli by first measuring the IPI between each syllable in the test stimulus. Each IPI between syllables was then multiplied by 2 and the IPI extended to this duration. All IPI manipulations were created using SoundEdit (version 2).

Amplitude

Justification Amplitude fluctuations in CLCs occur both at the level of each individual syllable and across the duration of the call. Although there does not seem to be a consistent amplitude envelope shape either within or between individuals, previous perceptual experiments suggest that this feature may be important for call recognition (Ghazanfar et al. 2002). In addition, the power of each syllable is not constant across the call, but rather shows an increase over the duration of the vocalization. The functional significance of both within- and between-syllable amplitude modulation for CLC is unknown. The CLCs used as stimuli in this condition consisted of the F0 and harmonics.

Between-syllable amplitude modulation In this condition, we manipulated the differences in acoustic energy that occur in the whistle syllables across CLCs. All stimuli in this condition were synthetic CLCs with a F0 and three upper harmonics. To determine the normal amplitude modulation between syllables in the CLC, we measured the RMS amplitude for each whistle for 50 CLCs (10 calls from 5 subjects). Results indicated that the mean RMS

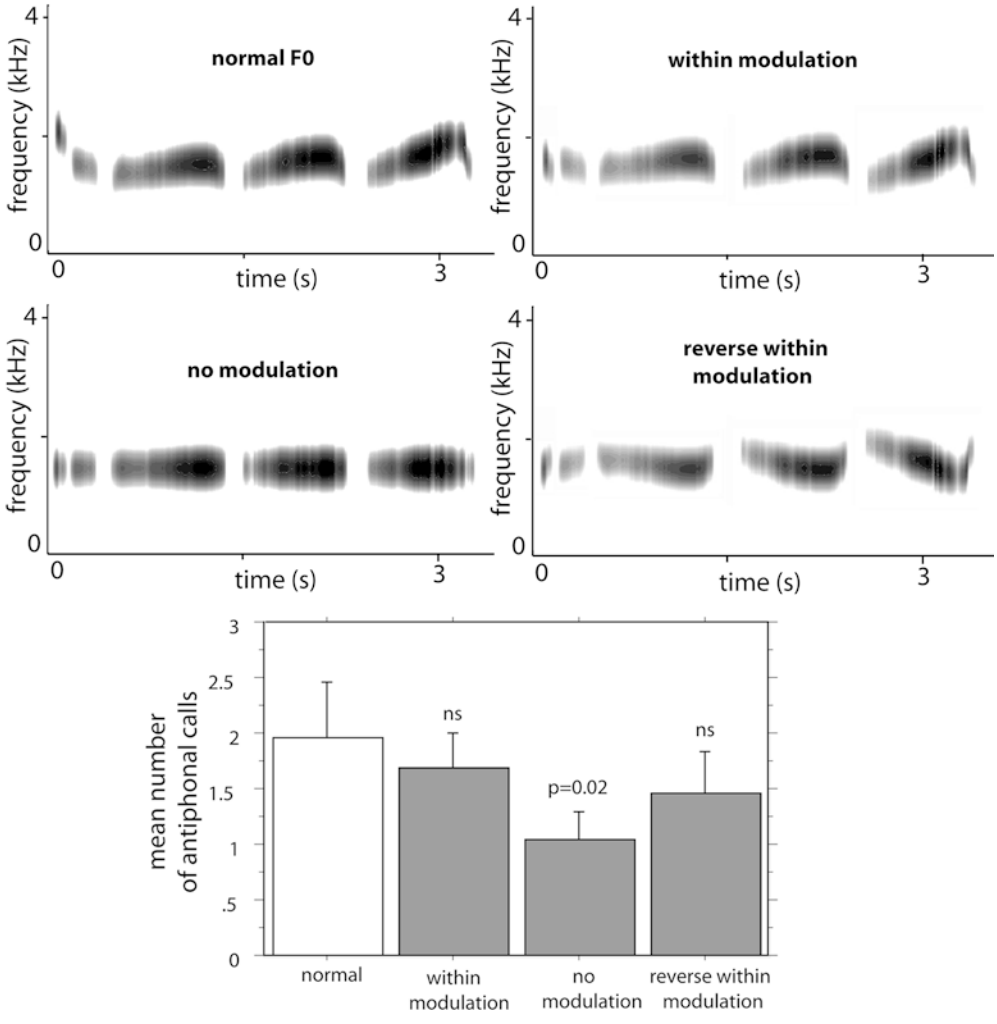

Fig. 3 Spectrograms of the four synthetic CLCs used as test stimuli in the within-frequency modulation test condition (experiment 2: condition 2) are shown at the *top*. *Below* is a bar graph showing the mean antiphonal calling responses for all stimuli with standard error bars. Antiphonal call responses to normal calls are shown *unshaded*, while responses to test stimuli are shown in *black*. For all stimulus types that elicited a statistically different antiphonal call rate from 'normal' calls, the *P* value is shown. (*ns*: not significantly different)

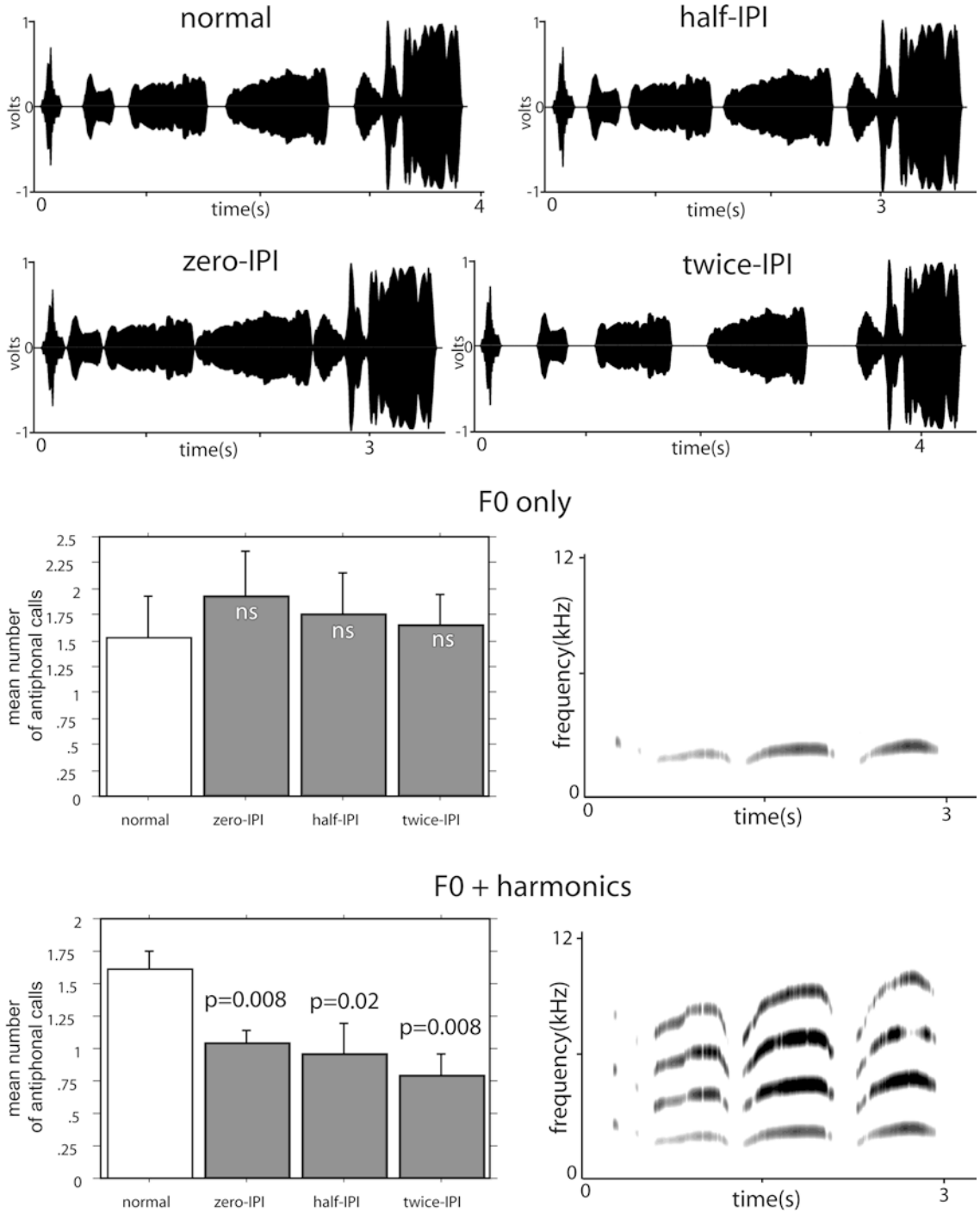

Fig. 4 Shown at the top are the amplitude envelopes of the four synthetic CLCs used in experiment 2: conditions 3 and 4 are shown at the *top*. Depicted *below* are bar graphs showing the mean antiphonal calling responses with standard error bars and spectrograms of the stimuli for the fundamental frequency (F0) only condition (condition 3) and F0 + 3 harmonics condition (condition 4). For all stimulus types that elicited a statistically different antiphonal call rate from 'normal' calls, the *P* value is shown. (*ns*: not significantly different, *IPI*: inter-pulse interval)

for whistle 1 = 2.8 μp, for whistle 2 = 4.5 μp and for whistle 3 = 3.9 μp. However, the range for each syllable was considerable. The maximum and minimum RMS amplitude for each syllable was: whistle 1: 0.26–7.45 μp; whistle 2: 0.28–8.22 μp; whistle 3: 0.46–9.01 μp. Overall, these measurements suggest that between-syllable amplitude modulation is variable, but generally increases over the duration of the call. To test the perceptual significance of this aspect of amplitude fluctuation we generated the following test stimuli. For the *ascending amplitude modulation* stimuli, we decreased the amplitudes of each whistle corresponding to ~1.5 μp for whistle 1, ~3.5 μp for whistle 2, and ~5.5 μp for whistle 3. We manipulated all whistles in the *no amplitude modulation* stimuli to have an RMS amplitude of ~3.5 μp. In the *descending amplitude modulation* condition, we manipulated whistles in the test stimuli to have the following RMS amplitude: whistle 1 = ~5.5 μp, whistle 2 = ~3.5 μp and whistle 3 = ~1.5 μp. Because of the variable nature of this feature, the RMS amplitude values for each whistle were chosen somewhat arbitrarily.

However, all values fall within the naturally produced range and represent an initial effort to determine the significance of this feature for call recognition. Although changing the power of each syllable also resulted in the biproduct of altering the amplitude slope within some syllables, the relative fine amplitude changes within each syllable were preserved.

Within-syllable amplitude modulation Here we manipulated the position of the amplitude peak within each syllable. Stimuli in this condition consisted of synthetic CLCs with a F0 and three harmonics. As the duration of each syllable in each CLC used as test stimuli was different, it was not possible to use a consistent slope in manipulating this structure. Rather the amplitude slope of each syllable was created by first determining the peak in the syllable from which the amplitude would ascend or descend and then created the slope that would allow a relatively constant change in amplitude. For the *peak-start* stimuli, the peak amplitude is positioned in the first 50 ms of each syllable. In the *middle-peak* stimuli, the peak amplitude is in the middle 50 ms of each syllable. The peak amplitude in the *end-peak* stimuli is in the final 50 ms of each syllable.

Analysis

Following previous studies (Ghazanfar et al. 2001, 2002, Miller et al. 2001), we defined antiphonal calling as the production of either a CLC or normal long call (Cleveland and Snowdon 1982) within 5 s of stimulus offset. As a subject's propensity to vocalize does vary due to factors beyond the control of the experimenter (e.g., weather, previous interactions with group mates), only test sessions in which subjects produced an antiphonal call to at least two test stimuli were used in the analysis. The number of antiphonal calls to each stimulus was calculated for each test session. Statistical differences in patterns of antiphonal calling were calculated using repeated-measure ANOVAs. As there were only two different test stimuli used in experiment 1 (natural versus synthetic calls) the *F*-score generated by these ANOVAs was a direct comparison between these stimulus types. However, in experiment 2 (signal manipulations) we presented subjects with four stimulus types in each test condition. To statistically test whether antiphonal calling in each of the three signal manipulations differed from 'normal' CLCs, we performed contrast analyses. As the *F*-score from a repeated-measures ANOVA is omnibus, contrast analyses are necessary for comparing responses between two individual test conditions. For the contrast analyses used here, subjects' antiphonal response to 'normal' CLCs was directly compared to their responses to each individual manipulated stimulus type.

Results

Experiment 1: natural versus synthetic calls

Condition 1: natural CLC versus natural F0

This condition tested whether the antiphonal calling rates differ between calls consisting of the natural harmonic structure (F0 plus 3–6 harmonics) and calls containing only the fundamental frequency (F0). The mean number of antiphonal calls to natural CLCs was 3.0 (SE = 0.45) and 2.5 (SE = 0.38) for natural F0 stimuli. However, this difference was not significantly different ($F_{(1,12)} = 3.37$, p = 0.09). In addition, there was no interaction between session number and stimulus type ($F(3,36) = 2.41$, p = 0.08) suggesting that the pattern of antiphonal calling was similar for all stimulus types across all test sessions.

Condition 2: natural F0 versus synthetic F0

This condition examined whether subjects show different patterns of antiphonal calling to natural and synthetic CLCs with only the F0 contour preserved. Results indicated that subjects antiphonally called somewhat more to natural F0 stimuli (mean = 3.03, SE = 0.49) than to Synthetic F0 calls (mean = 2.48, SE = 0.40), but this difference was not statistically significant [$F_{(1,7)} = 5.3$, $P = 0.06$]. Similarly, there was no effect of trial [$F_{(3,21)} = 1.88$, $P = 0.16$], and no interaction between these factors [$F_{(3,21)} = 1.01$, $P = 0.41$].

We noticed that subjects seemed to antiphonally call more to natural F0 calls than to synthetic F0 calls during the first two sessions, but at comparable rates during the second two trials. As a result, it was possible that subjects first treated synthetic and natural calls as different, but with increased familiarity perceived them as similar. Alternatively, it may be that subjects became familiar with the specific stimuli rather than synthetics in general. To test between these alternative explanations, we conducted a second series of trials using two new stimulus sets. The only difference was that subjects were only run on two test sessions, rather than four. We predicted that if subjects had become familiarized with synthetic calls, then there should be no significant difference in antiphonal calling rates between the two stimulus types. In contrast, if the familiarization was to the specific stimuli, then subjects should respond differentially to this new stimulus set. Data revealed that subjects showed no significant difference in antiphonal calling to either synthetic F0 calls (mean = 2.50, SE = 0.61) or natural F0 calls [mean = 3.21, SE = 0.57; $F_{(1,6)} = 2.17$, $P = 0.20$]. Similarly, there was no interaction between session number and stimulus type [$F_{(1,6)} = 1.24$, $P = 0.31$].

Condition 3: natural CLC (F0 + harmonics) versus synthetic CLC (F0 + harmonics)

Here we tested whether subjects antiphonally call to natural and synthetic CLCs consisting of a F0 and three harmonics at similar or different rates. The mean number of antiphonal calls to natural CLCs was 3.36 (SE = 0.61) and to synthetic CLCs was 3.57 (SE = 0.68). This difference in antiphonal calling was not statistically significant [$F_{(1,6)} = 0.52$, $P = 0.49$]. There was no main effect of test trial [$F_{(3,8)} = 0.32$, $P = 0.81$] and no interaction between stimulus type and test trial [$F_{(3,18)} = 0.69$, $P = 0.57$]. These results suggest that both natural and synthetic CLCs elicit statistically similar rates of antiphonal calling.

Experiment 2: signal manipulations

Frequency

Condition 1: between-syllable frequency modulation In this condition, we tested whether subjects antiphonally called at different rates to stimuli that varied in between-syllable frequency modulation. Data revealed differences in the mean number of antiphonal calls emitted in response to each stimulus type: normal, mean = 2.11, SE = 0.31; between-modulation, 1.61, SE = 0.18; no-modulation, mean = 1.11, SE = 0.25; reverse-modulation, 1.75, SE = 0.20. Analyses showed a significant effect of stimulus type [$F_{(3,18)} = 5.94$, $P = 0.005$; Fig. 2]. However, the same analysis showed no significant interaction between session number and stimulus type [$F_{(9,54)} = 2.01$, $P = 0.06$] suggesting that subjects responded to the stimuli in roughly the same manner across all sessions. We next conducted contrast analyses to determine whether the pattern of antiphonal responses for the unmanipulated synthetic call was significantly different from the three manipulated stimulus types. There were no differences in antiphonal calling for unmanipulated CLCs compared to between-modulation calls [$F_{(1,6)} = 3.0$, $P = 0.13$] and reverse-modulation calls [$F_{(1,6)} = 1.79$, $P = 0.24$], but a significant decrease in antiphonal calling for no-modulation calls [$F_{(1,6)} = 15.27$, $P = 0.008$].

Condition 2: within-syllable frequency modulation In this condition, we presented subjects with stimuli in which the between-syllable frequency modulation was held constant and the within-syllable frequency modulation manipulated. Data showed that antiphonal calling rates differed between stimulus types (normal, mean = 1.96, SE = 0.49; normal within-modulation, mean = 1.68, SE = 0.33; no-modulation, mean = 1.04, SE = 0.25; reverse within-modulation, mean = 1.46, SE = 0.37). Analyses revealed a significant main effect of stimulus type [$F_{(3,18)} = 4.75$, $P = 0.01$; Fig. 3], but no interaction between session number and stimulus type [$F_{(9,54)} = 0.99$, $P = 0.46$]. Contrast analyses revealed no significant differences in antiphonal calling between unmanipulated CLCs and both normal within-modulation calls [$F_{(1,6)} = 0.93$, $P = 0.37$] and reverse within-modulation calls [$F_{(1,6)} = 1.68$, $P = 0.24$]. However, results indicated a significant difference between unmanipulated CLCs and no-modulation calls [$F_{(1,6)} = 8.82$, $P = 0.02$].

Time

Condition 3: interpulse interval (synthetic F0) In this condition, we presented subjects with CLCs that varied in their IPI. For this condition, we eliminated all of the upper harmonics, preserving only the F0. Subjects' antiphonal call rates differed between the stimulus types (normal, mean = 1.54, SE = 0.40; zero-IPI, mean = 1.93, SE = 0.43; half-IPI, mean = 1.75, SE = 0.41; twice-IPI, mean = 1.64, SE = 0.30). Results indicated no difference in antiphonal response across the different stimulus types [$F_{(3,18)} = 1.06$, $P = 0.39$; Fig. 4] and no interaction between session number and stimulus type [$F_{(9,54)} = 1.04$, $P = 0.42$]. These analyses suggest that subjects' antiphonal response to all stimuli across all test sessions were statistically indistinguishable.

Condition 4: interpulse interval (F0 and harmonics) We presented subjects with stimuli manipulated in the same fashion as condition 5, but rather than present the F0 contour alone, we broadcast CLCs consisting of a F0 as well as three harmonics. Here subjects' pattern of antiphonal calling differed across the different stimulus types (normal, mean = 1.61, SE = 0.14; zero-IPI, mean = 1.04, SE = 0.10; half-IPI, mean = 0.96, SE = 0.23; twice-IPI, mean = 0.79, SE = 0.17). Analyses indicated a significant effect of stimulus type [$F_{(3,18)} = 8.24$, $P = 0.001$; Fig. 4] and no interaction between session number and stimulus type [$F_{(9,54)} = 0.80$, $P = 0.62$]. Contrast analyses comparing antiphonal calling to unmanipulated CLCs with the other stimulus types indicated that subjects antiphonally called at significantly lower rates to zero-IPI stimuli [$F_{(1,6)} = 16.34$, $P = 0.007$], half-IPI stimuli [$F_{(1,6)} = 8.76$, $P = 0.02$), and twice-IPI stimuli [$F_{(1,6)} = 16.53$, $P = 0.007$].

Amplitude

Condition 5: between-syllable amplitude modulation The goal of this condition was to test whether the amplitude modulation that arises between syllables is an important cue in signal recognition. The four stimulus types in this condition were the following: normal CLC, ascending modulation (amplitude increasing over successive whistles), no modulation (equal energy in each whistle syllable) and descending modulation (amplitude decreasing over successive whistles). The mean number of antiphonal calls produced in response to these stimulus types was as follows: normal, 1.64 (SE = 0.50); ascending modulation, 1.32 (SE = 0.34); no modulation, 1.43 (SE = 0.33); descending modulation, 1.46 (SE = 0.33). There was no significant difference in antiphonal calling across the different stimuli [$F_{(3,18)} = 1.83$, $P = 0.18$; Fig. 5]. Further, no interaction was observed between stimulus type and test session [$F_{(9,54)} = 1.32$, $P = 0.25$] suggesting that the pattern of antiphonal calling was consistent across test sessions.

Condition 6: within-syllable amplitude modulation In this condition, we presented subjects with stimuli that varied in the amplitude modulation within syllables. For each manipulated CLC, the peak modulation was manipulated to occur either at the start (first 50 ms), middle (middle 50 ms), or end (final 50 ms) of each syllable. Data showed some differences in the number of

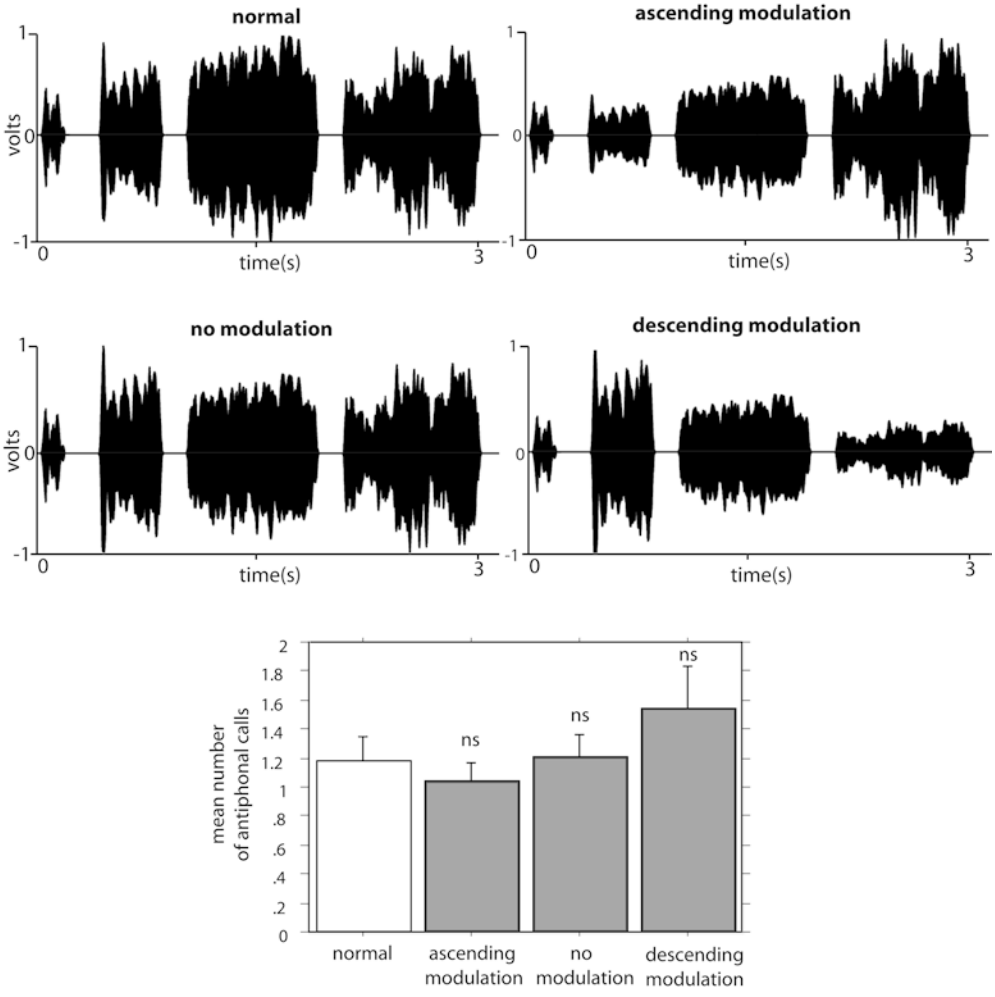

Fig. 5 The amplitude waveforms of the four synthetic stimuli used in experiment 2: condition 5 are shown at the *top*. *Below* is a bar graph showing the mean antiphonal calling responses for all stimuli with standard error bars. Antiphonal call responses to normal calls are shown *unshaded*, while responses to test stimuli are shown in *black*. For all stimulus types that elicited a statistically different antiphonal call rate from 'normal' calls, the *P* value is shown. (*ns*: not significantly different)

antiphonal calls produced in response to each of these stimuli (normal: mean = 1.18, SE = 0.17; start-peak: mean = 1.04, SE = 0.13; middle-peak: mean = 1.21, SE = 0.15; end-peak: mean = 1.54, SE 0.30). However, analyses revealed no significant difference in antiphonal calling across the different stimulus types [$F_{(3,18)} = 0.40$, $P = 0.76$; Fig. 6]. As such, we did not perform additional contrast analyses. Additionally, there was no interaction between stimulus type and test session [$F_{(9,54)} = 0.79$, $P = 0.62$] suggesting that subjects' pattern of antiphonal response was statistically similar across all sessions.

Discussion

The goal of our experiments was to examine the role of frequency, time and amplitude for recognition of the cotton-top tamarin CLC. Following previous studies (Ghazanfar et al. 2001, 2002; Miller et al. 2001), we used the tamarins' natural, species-typical antiphonal calling response to hearing CLCs as the primary measure of call recognition. We hypothesized that if experimental changes in acoustic morphology cause a significant decrease in antiphonal responses then the targeted feature is important for signal recognition, at least in our testing environment. In contrast, if subjects respond at similar rates to normal and manipulated calls, then we the feature is presumed to be irrelevant for recognition. Note that the issue is not whether tamarins might be able to discriminate manipulated calls under other conditions, such as during psychophysical tests using operant training. Rather, our focus is on their natural, untrained, responses to calls, and whether the information encoded is sufficient to elicit such responses. In this sense, we follow the distinction most clearly articulated by Nelson (1988) between just noticeable differences (JNDs) and just meaningful differences (JMDs). Our results refer to JMDs, and not necessarily to JNDs.

Studies of anurans and birds have often used synthetic stimuli in studies of vocal signal recognition because the signals can be readily manipulated across multiple dimensions (Nelson 1988; Gerhardt 1991; Ryan and Rand 1995). To test whether tamarins antiphonally called at similar rates to both synthetic CLCs and naturally produced exemplars of this call type, we presented subjects with exemplars of natural CLCs and synthetic replicas of these calls in experiment 1. As this study involved the use of synthetic CLCs consisting of either a F0 and multiple harmonics or the F0 alone, we conducted comparisons for both of these stimuli with their

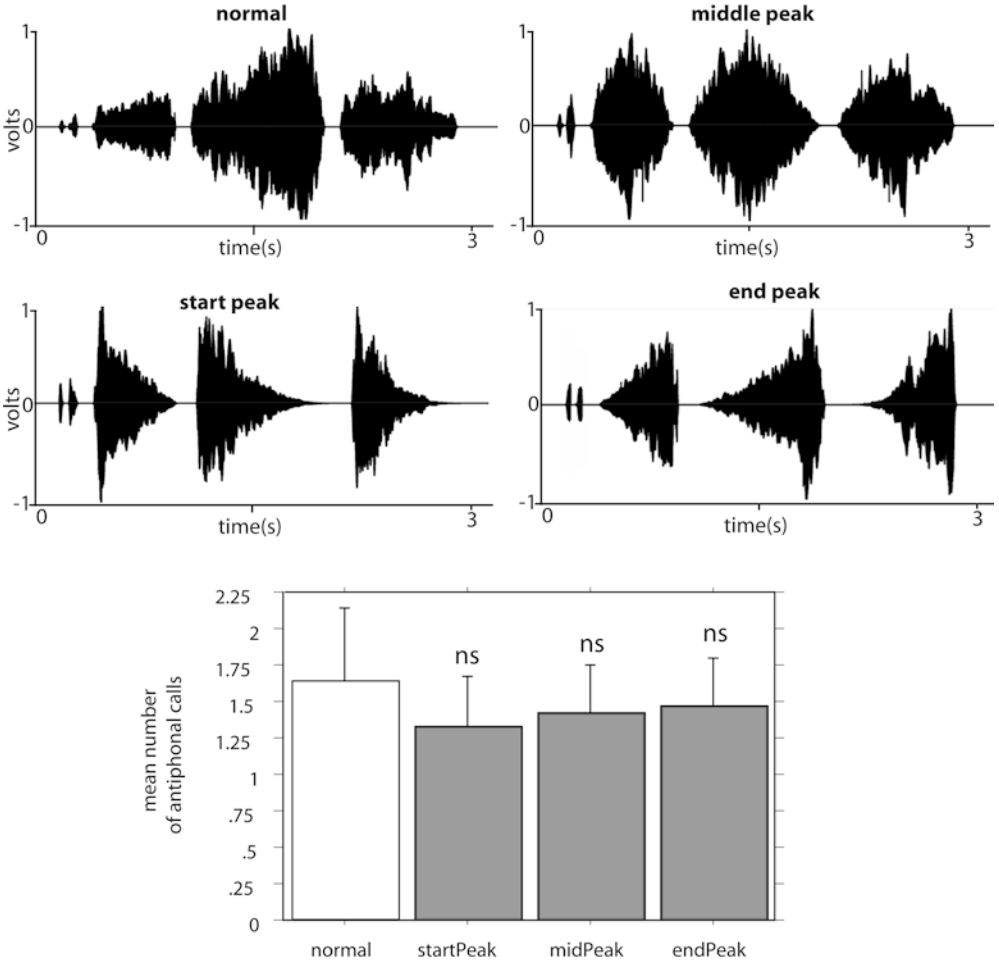

Fig. 6 The amplitude waveforms of the four synthetic test stimuli used in experiment 2: condition 6 are shown at the *top*. *Below* is a bar graph showing the mean antiphonal calling responses for all stimuli with standard error bars. Antiphonal call responses to normal calls are shown *unshaded*, while responses to test stimuli are shown in *black*. For all stimulus types that elicited a statistically different antiphonal call rate from 'normal' calls, the *P* value is shown. (*ns*: not significantly different)

respective naturally produced CLC. Results revealed no statistical difference in antiphonal calling to synthetic and naturally produced CLCs (Table 1). It should be noted that although both natural and synthetic calls elicit comparable rates of antiphonal calling, the tamarins may not perceive synthetic calls as being identical to natural calls. However, data suggest that tamarins perceive synthetic calls to be sufficiently similar to natural calls so as to elicit a statistically similar rate of antiphonal calls. Therefore, we are able to use synthetic signals to examine the relevance of particular acoustic features in CLCs by manipulating the synthetic signal along thee acoustic dimensions: frequency, time and amplitude.

Table 1 Summary of results

Condition		Significantly different calling rate?
Experiment 1		
1	Natural CLC versus naturol F0	No
2	Natural F0 versus synthetic F0	No
3	Natural CLC versus synthetic CLC	No
Experiment 2		
1	Between-syllable frequency modulation	Yes
2	Within-syllable frequency modulation	Yes
3	IPI (F0 only)	No
4	IPI (F0 + harmonics)	Yes
5	Between-syllable amplitude modulation	No
6	Within-syllable amplitude modulation	No

CLC combination long-call, *F0* fundamental frequency, *IPI* interpulse interval

Many species of animals emit calls with considerable modulation in the frequency contour (Hauser 1996). Evidence suggests that changes in frequency contour provide reliable cues to species and call identity (Green 1975; Mitani and Marler 1989; Winter 1969). Acoustic analyses show that the CLC contains frequency contour modulation both within and between each syllable (Weiss et al. 2001). To test whether frequency modulation was relevant for CLC recognition, we presented tamarins with CLCs in which we manipulated the modulation of the frequency contour. Results indicated that modulation both between- (experiment 2: condition 1, Fig. 2) and within-syllables (experiment 2: condition 2, Fig. 3) was important for recognition of the CLC. In both conditions, however, subjects antiphonally called at similar rates to stimuli in which the frequency modulation was normal and stimuli in which the modulation was reversed. This suggests that the underlying recognition mechanisms may be tuned to detect changes in the frequency contour, but are not specific in terms of the direction of the modulation.

Previous studies by Ghazanfar et al. (2001, 2002) suggested that the temporal pattern of the CLC, specifically the IPI, was an important feature for signal recognition in antiphonal calling. Building on these studies, we conducted two test conditions in which we manipulated the IPI of test CLCs. In experiment 2 (condition 3) we presented subjects with stimuli consisting of only the fundamental frequency. Here subjects showed no difference in antiphonal calling between any of the test exemplars (Fig. 4). In experiment 2 (condition 4), however, we presented subjects with stimuli consisting of the fundamental frequency and three harmonics. Results from this condition revealed that subjects antiphonally called significantly less to all manipulated CLCs compared to normal CLCs. This pattern of behavior may provide insights into how the tamarin auditory system processes CLCs to encode relevant temporal information. Specifically, while the frequency contour can be assessed from the fundamental frequency, the temporal pattern of the signal is obtained from the upper harmonics. This pattern of recognition may occur because higher frequency components of a signal attenuate in rainforests—the tamarins' natural habitat—at a faster rate then lower frequency components (Brown 2003). Further, IPI information attenuates quickly in vocalizations emitted in forests because of high levels of ambient noise that interfere with the efficacy of the signal (Waser and Waser 1977). As a result, only vocalizations with little environmental noise interference are likely to have reliable IPI information available to signal receivers. As the tamarin auditory recognition system attempts to identify a signal as a conspecific, a simple heuristic may be to ignore IPI information if the upper harmonics are degraded or missing because it is likely that IPI will not be a reliable cue.

In the final set of test conditions, we explored the relevance of amplitude fluctuation for recognition of the CLC. We manipulated amplitude at the individual syllable level by modifying the shape of the amplitude envelope (experiment 2; condition 5), as well as manipulating the amplitude across the whole call by changing the relative energy of each syllable (experiment 2; condition 6). Results indicated that subjects antiphonally called at comparable rates to both normal and manipulated CLCs in both test conditions suggesting that amplitude changes, both within and between syllables, are not significant features for CLC recognition. This result is not surprising given that CLCs serve as long-range signals and transmit over long distances in their natural forest habitat. Studies of the habitat acoustics show that tropical rainforests severely perturb the amplitude structure of vocalizations making it an unreliable cue for call recognition (Wiley and Richards 1978; Waser and Brown 1986).

Results from these experiments suggest that tamarins do not rely on acoustic features within a single dimension for recognition of their species-specific CLC, but rather use a combination of frequency and temporal cues (Table 1). This suggests that the tamarin auditory recognition system is flexible, processing and recognizing signals with either perturbed or missing information. This conclusion is consistent with two previous studies of the tamarin's antiphonal calling behavior. First, Miller et al. (2001) showed that tamarins are able to identify the 'whistles' in the CLC when the middle portion of the syllable is replaced by white noise; this suggests that tamarins are able to recognize components of the CLC when considerable acoustic information is entirely missing. Second, Ghazanfar et al. (2002) found that tamarins produce antiphonal calls at similar rates to normal CLCs and CLCs in which the spectral content is replaced by white noise, but the species-typical amplitude envelope and IPI are preserved. This suggests that if significant acoustic information is missing, tamarins can use alternative acoustic features to recognize CLCs in some contexts. Overall, the pattern of results emerging from studies of antiphonal calling in tamarins suggests that this behavior reflects the environment in which it evolved: a dense rainforest with high levels of ambient noise, an acoustic environment known to cause rapid signal degradation (Waser 1977; Brown et al. 1995). Given the high degree of signal perturbation that arises for long-distance vocalizations, it would be adaptive for the tamarin auditory recognition system to be able to recognize degraded conspecific vocalizations. Thus, evolving a system that uses multiple cues for signal recognition would facilitate recognizing conspecific vocalizations in a noisy environment. To further explore this issue, we have begun a study of the transmission properties of the CLC in different habitats, including the tamarins' species typical rainforest (C.T. Miller and Palleroni, unpublished data).

Neuroethological studies of communication in songbirds, anurans and insects have yielded an abundance of data on the neural mechanisms that underlie communicative behavior in animals. At present, little is known about this process in primates (Newman and Wollberg 1973a; Winter and Funkenstein 1973; Bieser 1998). The antiphonal calling behavior of cotton-top tamarins, and other primates, represents a potential assay to investigate the neural mechanisms of primate communication. At the behavioral level, antiphonal calling is highly stereotyped, easily elicited and occurs at a frequent rate. Further, results from this study and others (Ghazanfar et al. 2001, 2002; Miller et al., 2001) show that the antiphonal calling paradigm can be used as a behaviorally relevant method to assess signal recognition. Recent neurophysiological studies of the closely related common marmoset (*Callithrix jacchus*), a species that also engages in antiphonal calling (C.T. Miller et al., unpublished data), suggest that single-unit recordings of neural activity are possible while animals both hear conspecific vocalizations (Wang and Kadia 2001) and while they are produced (Eliades and Wang 2003). At present no studies have simultaneously recorded from vocal perception and production areas in primate cortex to investigate this sensory-motor interaction, but

employing the antiphonal calling paradigm may provide a neuroethological approach to such an investigation.

Acknowledgements We would like to thank Matt Kamen, Teddy Jones, and Grace Wang for their invaluable help running these experiments. Asif Ghazanfar, Roian Egnor and two anonymous reviewers provided helpful comments on an earlier draft that greatly improved the manuscript. We thank the NERPRC for providing some of the initial founding members of the colony. This work was supported by a NIH Individual NRSA Predoctoral Fellowship to C.T.M. and grants from the NSF, NIH and McDonnell Foundation to M.D.H.

References

Beeman K (2001) SIGNAL user's guide. Engineering Design, Belmont, MA

Bieser A (1998) Processing of twitter-call fundamental frequencies in insula and auditory cortex of squirrel monkeys. Exp Brain Res 122:139–148

Brown CH (2003) Ecological and physiological constraints for primate communication. In: Ghazanfar AA (ed) Primate audition: ethology and neurobiology. CRC Press, New York, pp 127–149

Brown CH, Gomez R, Waser PM (1995) Old world monkey vocalizations: adaptations to the local habitat? Anim Behav 50:945–961

Cleveland J, Snowdon CT (1982) The complex vocal repertoire of the adult cotton-top tamarin, *Saguinus oedipus oedipus*. Z Tierpsychol 58:231–270

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

Doupe AJ, Solis MM (1999) Song- and order-selective auditory responses emerge in neurons of the songbird anterior forebrain during vocal learning. In: Hauser MD, Konish M (eds) The design of animal communication. MIT Press, Cambridge, MA, pp 343–368

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

Emlen ST (1972) An experimental analysis of the parameters of bird song eliciting species recognition. Behaviour 41:130–171

Gerhardt HC (1991) Female mate choice in treefrogs: static and dynamic acoustic criteria. Anim Behav 42:615–636

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans. Chicago University Press, Chicago

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

Ghazanfar AA, Smith-Rohrberg DL, Pollen A, Hauser MD (2002) Temporal cues in the antiphonal calling behavior 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

Green S (1975) Dialects in Japanese monkeys: vocal learning and cultural transmission of locale-specific behavior? Z Tierpsychol 38:304–314

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

Hauser MD, Fowler C (1992) Declination in fundamental frequency is not unique to human speech: evidence from nonhuman primates. J Acoust Soc Am 91:363–369

Konishi M (1985) Birdsong: from behavior to neuron. Annu Rev Neurosci 8:125–170

Konishi M (1994) An outline of recent advances in birdsong neurobiology. Brain Behav Evol 44:279–285

Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. J Comp Physiol Psychol 71:1–25

Marler P (1991) Song learning behavior: the interface with neuroethology. TINS 14:199–206

Marler P (1997) Three models of song learning: evidence from behavior. J Neurobiol 33:1-16

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

Miller CT, Weiss DJ, Hauser MD (2003a) Mechanisms of acoustic perception in cotton-top tamarins. In: Ghazanfar AA (ed) Primate audition: behavior and neurobiology. CRC Press, Boca Raton, pp 43–60

Miller CT, Flusberg S, Hauser MD (2003b) Interruptibility of cotton-top tamarin long calls: implications for vocal control. J Exp Biol 206:2629–2639

Mitani J, Marler P (1989) A phonological analysis of male gibbon singing behavior. Behaviour 109:20–45

Moody DB, Stebbins WC (1989) Salience of frequency modulation in primate communication. In: Dooling RJ, Hulse SH (eds) The comparative psychology of audition. Lawrence Erlbaum, Hillsdale, NJ, pp 353–378

Nelson DA (1988) Feature weighting in species song recognition by the field sparrow (*Spizella pusilla*). Behaviour 106:158–182

Nelson DA (1997) Social interaction and sensitive phases for song learning: a critical review. In: Snowdon CT, Hausberger M (eds) Social influences on vocal development. Cambridge University Press, Cambridge, UK

Nelson DA, Marler P (1990) The perception of birdsong and an ecological concept of signal space. In: Stebbins WC, Berkley MA (eds) Comparative perception, vol II: complex signals. Wiley, New York, pp 443–478

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

Newman JD, Wollberg Z (1973b) Responses of single neurons in the auditory cortex of squirrel monkeys to variants of a single call type. Exp Neurol 40:821–824

Owren MJ (1990) Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans. II. Synthetic calls. J Comp Psychol 104:29–40

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

Ryan MJ (2001) Anuran communication. Smithsonian Institution Press, Washington, D.C.

Ryan MJ, Rand AS (1995) Female responses to ancestral advertisement calls in Tungara frogs. Science 269:390–392

Searcy WA, Brenowitz EA (1988) Sexual differences in species recognition of avian song. Nature 332:152-154

Searcy WA, Podos J, Peters S, Nowicki S (1995) Discrimination of song types and variants in song sparrows. Anim Behav 49:1219–1226

Sen K, Theunissen FE, Doupe AJ (2001) Feature analysis of natural sounds in the songbird auditory cortex. J Neurophysiol 86:1145–1458

Snowdon CT (1993) A vocal taxonomy of the callitrichids. In: Rylands AB (ed) Marmosets and tamarins: systematics, behaviour and ecology. Oxford University Press, New York

Soha JA, Marler P (2000) A species-specific acoustic cue for selective song learning in the white-crowned sparrow. Anim Behav 60:297–306

Solis MM, Doupe AJ (1997) Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. J Neurosci 17:6447–6462

Solis MM, Doupe AJ (2000) Compromised neural selectivity for song in birds with impaired sensorimotor learning. Neuron 25:109–121

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

Waser PM (1977) Experimental playbacks show vocal mediation of avoidance on a forest monkey. Nature 255:56–58

Waser PM, Brown CH (1986) Habitat acoustics and primate communication. Am J Primatol 10:135–154

Waser PM, Waser MS (1977) Experimental studies of primate vocalization—specialization for long distance propagation. Z Tierpsychol 43:239–263

Weiss DJ, Hauser MD (2002) Perception of harmonics in the combination long call of cotton-top tamarins (*Saguinus oedipus*). Anim Behav 64:415–426

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 Psychol 11:258–271

Whaling CS, Solis MM, Doupe AJ, Soha JA, Marler P (1997) Acoustic and neural bases for innate recognition of song. PNAS 94:12694–12698

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. Behav Ecol Sociobiol 3: 69–94

Winter P (1969) Dialects in squirrel monkeys: vocalizations of the Roman arch type. Folia Primatol 10:216–229

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