Temporal cues in the antiphonal long-calling behaviour of cottontop tamarins

ASIF A. GHAZANFAR, DUNCAN SMITH-ROHRBERG, ALEX A. POLLEN & MARC D. HAUSER

Primate Cognitive Neuroscience Laboratory, Department of Psychology, Harvard University

In primates, long-range communication is often mediated by the use of 'long' (or 'loud') calls. Beyond the acoustic classification of these calls and descriptions of the behavioural context in which they are produced, few experimental studies have examined how species-typical information is encoded in the structure of these signals. We present the results of eight experimental conditions designed to isolate the mechanisms underlying the perception of long calls in cottontop tamarins, *Saguinus oedipus*. Our procedure involved presenting a combination of naturally produced and experimentally manipulated long calls to individuals isolated from their group, and then recording the relationship between signal design and the production of antiphonal long calls by the test subject. Tamarins did not distinguish between normal calls and time-reversed or pitch-shifted long calls, but normal response rates did require the species-typical amplitude envelope. Furthermore, there was suggestive evidence that the number of syllables and the syllable rate may also influence antiphonal calling responses. We discuss these results in terms of the mechanisms of vocal recognition in primates and other taxa.

© 2002 The Association for the Study of Animal Behaviour. Published by Elsevier Science Ltd. All rights reserved.

Many species have evolved signals that can transmit information to conspecifics over long distances. Because of their conspicuousness, these signals have been the focus of much research including detailed studies of insects (Ewing 1989), frogs (Fritzsch et al. 1988) and birds (Kroodsma & Miller 1996). Ethologists studying vocal communication in these taxa (birds: e.g. Emlen 1972; Nelson 1988; Holland et al. 2000; Lengagne et al. 2000; anurans: e.g. Narins & Capranica 1978; Ryan & Rand 1990; Gerhardt 1991) have explored the perceptual mechanisms by manipulating features of conspecific signals that are hypothesized to be relevant and then measuring species-typical responses of subjects to playbacks of these manipulated signals.

In primates, long-range communication is often mediated by the use of 'long' or 'loud' calls. Primates generally produce long calls in the context of territorial encounters, mate attraction and isolation/group cohesion (Marler 1968; Waser 1982; Mitani 1985). In many species, these calls often elicit similar 'antiphonal' calls from conspecifics (Waser 1977; Cleveland & Snowdon 1982; Ghazanfar et al. 2001), whereas in other species anti-

Correspondence and present address: A. A. Ghazanfar, Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, Tuebingen 72076, Germany (email: asifg@tuebingen.mpg.de). D. Smith-Rohrberg, A. A. Pollen and M. D. Hauser are at the Primate Cognitive Neuroscience Laboratory, Department of Psychology, Harvard University, Cambridge, MA 02138, U.S.A.

phonal calls are conspicuously absent (Cheney et al. 1996; Rendall et al. 2000). Acoustically, long calls consist of sequences of similar acoustic units and/or sequences of different-sounding acoustic units and serve as location signals for conspecifics (Waser 1977; Newman et al. 1978; Robinson 1979; Mitani 1985; see Miller & Ghazanfar 2002 for a review). Beyond the acoustic classification of these calls and descriptions of the behavioural context in which they are produced, few experimental studies have examined how species-typical information is encoded in the structure of these primate vocal signals (Robinson 1979; Mitani & Marler 1989; Ghazanfar et al. 2001). As a result, while the function of primate long calls has received much attention (Boinski 2000), the perceptual mechanisms underlying primate long-calling behaviour have been rarely studied.

The vocal behaviour of the cottontop tamarin, *Saguinus oedipus* (hereafter, tamarins), a small, arboreal New World primate species, is particularly well suited to investigations of the perceptual mechanisms underlying long calling. First, in response to social isolation, both wild and captive tamarins will readily produce a long call (Andrew 1963; Neyman 1978; Cleveland & Snowdon 1982). When conspecifics hear such calls, they will often respond with their own antiphonal long calls (Cleveland & Snowdon, 1982; Ghazanfar et al. 2001). Second, we know much about their vocal interactions and long-calling behaviour in different social contexts (Snowdon

et al. 1983; McConnell & Snowdon 1986; Ghazanfar et al. 2001). Finally, extensive acoustic analyses of the tamarin long call reveal that individual, sexual and group identities are encoded in the structure of the long call, and individual recognition is supported by playback experiments (Weiss et al. 2001).

According to acoustic and behavioural criteria, there are three long-call types: the normal long call, the quiet long call, and the combination long call (CLC; Cleveland & Snowdon 1982). During social isolation, tamarins produce CLCs. These calls consist of one or more 'chirps' followed by one or more 'whistles' (Cleveland & Snowdon 1982; Weiss et al. 2001). Our previous study indicated that socially isolated tamarins are more likely to call antiphonally, and will call more frequently, following playbacks of CLCs than to individual chirps or whistles (Ghazanfar et al. 2001). Thus, there are acoustic features embedded in the structure of the CLC that elicit antiphonal calls from conspecifics. In the present study, we investigated the acoustic features of CLCs that are necessary to elicit antiphonal responses in socially isolated tamarin receivers.

METHODS

We tested 12 adult tamarins. They were members of a colony consisting of 13 adults, two subadults and four infants. All adults were born in captivity at the New England Regional Primate Center, Southborough, MA, U.S.A. and then housed at the Primate Cognitive Neuroscience Laboratory at Harvard University, where the subadults and infants were born. Subjects were housed in groups of two or three adults and their offspring in a home cage (1.8 × 1.5 m and 0.76 m high) made of steel wire and Plexiglas. Each home cage contained nestboxes, tree branches and other forms of environmental enrichment. In all, the colony consisted of five mated pairs and three family groups. They were kept on a 12:12 h light: dark cycle and were fed once a day in the early evening. Their diet consisted of Purina tamarin and marmoset chow, supplemental vitamins and sunflower seeds and this diet was supplemented by food received during experiments (typically raisins, occasionally Froot Loops and marshmallows). Subjects had ad libitum access to water.

We tested tamarins, over a period of 9 months (June 2000 to February 2001), for their responsiveness to playbacks of sequences of CLCs alternated with manipulated exemplars of CLCs. During the experiments, two subjects (one male, one female) had high long-calling rates which precluded their participation in our study. These individuals simply called at such high rates that it was impossible to discern whether they were calling antiphonally or simply long calling spontaneously according to our criteria. In addition, one male did not long call (spontaneous or otherwise) when placed in our experimental apparatus. One female completed one experimental condition (the intersyllable interval contraction) before she died. Thus, for all but one condition reported below, there were eight subjects, three males and five females.

Apparatus

During these playback experiments, the tamarins were placed in a wire-and-cloth test cage (45 × 45 cm and 20 cm high). This cage was housed in an acoustic chamber (Model 400-A, Industrial Acoustics Company, Inc., New York, U.S.A.). A thin black cotton sheet hung behind the cage which concealed an Alesis Monitor One speaker (frequency range, 45–18 000 Hz) mounted on a shelf above the box, behind and to the left. An Alesis RA-100 amplifier drove the speaker. This system has been used in previous studies of tamarin perception (Ghazanfar et al. 2001; Weiss et al. 2001).

We used a video camera to monitor the sessions outside the acoustic chamber. Playbacks of calls were controlled with a Hypercard program (written by W. T. Fitch) on a Power Macintosh 7100/80 AV. Calls were played through an Audiomedia II sound card outputted to the speaker. Experimenters entered data and kept track of time elapsed between trials using the Hypercard program.

Stimuli

Only CLCs, and manipulated variants of CLCs, were used as stimuli in the present study. All calls were five-syllable calls (containing one or two chirps followed by three or four whistles). We recorded spontaneously produced CLCs, of socially isolated tamarins, with a Tascam DAT (DP-1) and a Sennheiser ME-60 microphone. The calls were digitized and hi-pass filtered using Sound Designer II software. To determine the cut-off frequency for the filter, we examined the spectrogram (1024 FFT) of each call to determine the lowest frequency. After filtering the call below that frequency, we normalized all of the calls to 100%. We screened each call by examining the spectrograms and listening to each exemplar in the playback chamber to ensure that they were free of any artefacts. Only calls judged to be of high recording quality were selected for use in this experiment.

We selected 10 CLC exemplars from each of 10 different adults (five males and five females) from our colony, yielding a total of 100 calls. We then generated five unique call sets of 20 calls (two exemplars per caller), and three more call sets made up of exemplars picked at random from the other five call sets. Call sets/playback sequences were organized in the following manner. In each call set of 20 calls, one exemplar from a caller was normal and the other exemplar from the same individual was manipulated acoustically. The normal and manipulated exemplars were then switched to make an alternative version of the call set (e.g. Call set 1A and Call set 1B). Consequently, subjects heard both normal and manipulated versions of every CLC exemplar (Fig. 1).

By using subjects from the same colony as the source for our call sets, we were able to collect a large number of excellent-quality long-call exemplars with which we could make multiple call sets, and a unique set of calls was available for almost every condition. One potential drawback is that the calls were from tamarins familiar to each of the subjects. Since individual identity is encoded in the structure of long calls and there is evidence
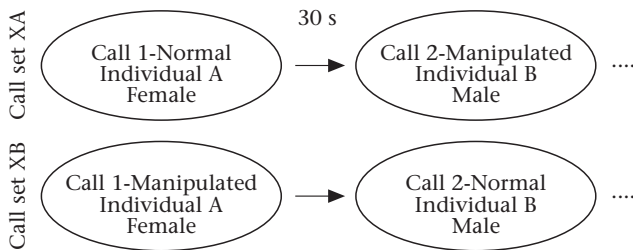
Figure 1. Structure of call sequences used in playback experiments. During an experiment, normal five-syllable combination long calls (CLCs) were alternated with manipulated versions of CLCs. There were 20 playback stimuli in each call set (two exemplars each from 10 different tamarins), 10 normal and 10 manipulated. Playbacks of each stimuli were separated by a minimum of 30 s. Subjects always heard both the normal calls and the alternative stimuli from each of the 10 individuals during an experiment. Exemplars from the same individual were spaced temporally as far apart as possible in the call sequence. Call sequence A was presented in the reverse order as call sequence B (e.g. the manipulated call preceded the normal call).

for recognition of individuals from habituation–discrimination experiments (Weiss et al. 2001), it is possible that this affected our pattern of results. For example, subjects could respond preferentially to his/her cagemate's long calls, or suppress antiphonal calls upon hearing 'self-calls'. However, several considerations suggest that this was not the case. Weiss et al. (2001) found no evidence that tamarins were more or less likely to discriminate their own calls from another individual than they were to discriminate two different individuals. In a separate study, A. A. Ghazanfar, J. I. Flombaum & M. D. Hauser (unpublished data) used the antiphonal calling method presented here to show that tamarins do not preferentially call to (1) self-calls compared to other familiar individuals within the colony (see also Snowdon et al. 1983 for similar findings), (2) cagemate calls compared to other familiar individuals, and (3) familiar versus foreign individuals. We therefore conclude that using exemplars from the same colony does not present any problems for our experiments.

Signal Manipulations

As there are only a handful of playback studies on the perceptual mechanisms underlying primate call recognition (Fischer 1998; Hauser et al. 1998; Fischer et al. 2001; Ghazanfar et al. 2001), our choice of signal manipulations was motivated entirely by the studies of long-distance communication in other taxa (Emlen 1972; Gerhardt 1981; Nelson 1988; Ryan & Rand 1990). Each manipulation was done on a call-by-call basis. Exemplars of each signal manipulation are presented in the Results. The following signal manipulations were made using SoundForge 4.5 (Sonic Foundry, Inc., Madison, Wisconsin, U.S.A.), CoolEditPro 1.1 (Syntrillium Software, Phoenix, Arizona, U.S.A.) or Signal (Engineering Design, Belmont, Massachusetts, U.S.A.).

(1) Local reversed. In this manipulation, the overall order of long call syllables was normal but individual syllables were reversed. This changed the local spectral contour but preserved the overall global spectral contour (i.e. overall, the calls had an upward frequency sweep).

(2) Global reversed. The entire call was temporally reversed. These stimuli retained all the spectral information, but the syllables and the temporal order of syllables were reversed. In normal CLCs, there is usually an upward frequency sweep. This manipulation reversed the spectral contour of the signal so that, overall, there was a downward frequency sweep.

(3) Low-pitch shift. The frequency of calls carries information (e.g. body size) which receivers may use in their decision to respond to a signal (Ryan & Wilczynski 1988; Fitch 1997; Fitch & Kelley 2000; Bosch et al. 2000). We lowered the pitch of long calls by reducing the fundamental frequency of each call by one-half. Thus, for example, a call whose fundamental frequency was originally 2000 Hz was lowered to 1000 Hz. The average fundamental frequency ± SD for long calls (measured across all syllables) was 2395 ± 852 Hz (B. Garibaldi, D. Weiss & M. D. Hauser, unpublished data). On average, therefore, our manipulation generated a call that was 1.4 SDs away from the mean fundamental frequency. As pitch shifting was done on a call-by-call basis, some manipulated calls were within the species-typical range while others were beyond it.

(4) High-pitch shift. The fundamental frequency of the long call was doubled. For example, a fundamental frequency of 2000 Hz was shifted to 4000 Hz. On average, this created a high-pitched call that was 2.8 SDs from the mean fundamental frequency.

(5) Amplitude-modulated white noise. In this condition, the amplitude envelope of each individual CLC was applied to white noise. This generated stimuli that retained the species-typical temporal features of CLCs, but the species-typical spectral content was replaced by a random spectrum with a bandwidth of 50–20 000 Hz. These stimuli were created in Signal by first extracting the amplitude envelope of the normal call (decay constant= 5 ms) using the envelope command. The amplitude envelope was then multiplied by an appropriate duration of white noise.

(6) Gated white noise. In these stimuli, both the species-typical spectral content and the fine structure of the species-typical amplitude envelope were replaced by white noise and a square-pulsed function. These stimuli were created in Signal using the gate function, which captures the on–off temporal pattern of syllables but removes the fine temporal structure of amplitude modulation within the syllables. The gated signal was then multiplied by an appropriate duration of white noise, producing stimuli consisting of bursts of white noise of the same overall duration and intersyllable intervals as the normal CLCs.

(7) Intersyllable interval contraction. The silent intervals between syllables in the CLCs were edited out. In our colony of tamarins, these intersyllable intervals averaged 119.95 ms. On average, therefore, this manipulation reduced the overall duration of CLCs from 2542.4 ms to 2062.1 ms. In such stimuli, the syllable rate increased from 2.0 to 2.4 syllables/s. All other acoustic features were the same as in normal CLCs.

(8) Intersyllable interval expansion. In this condition, blocks of silence were inserted in the intersyllable intervals to expand this feature to twice the maximum length (668 ms) observed in our colony (B. Garibaldi, D. Weiss & M. D. Hauser, unpublished data). This extended the average duration of CLCs from 2542.4 ms to 4734.2 ms and decreased the syllable rate from 2.0 to 1.1 syllables/s.

In the intersyllable interval expansion condition, we noted during the playback of the first call set that subjects were interrupting the expanded calls. In other words, subjects were calling antiphonally before the end of the stimulus. For all subsequent runs of this condition, therefore, we recorded the entire session on a DAT. Vocal responses were then acquired online using Signal's real-time acquisition module and then analysed to determine (1) how many normal and expanded calls were interrupted; and (2) during/following which syllable the interruptions occurred. For this latter analysis, interruptions were scored according to which syllable (1–5) preceded, or was disrupted by, the antiphonal call from the subject (see Results).

Experimental Design and Playback Procedure

Subjects were tested between 0700 and 1000 hours prior to their morning feeding. To begin a session, we placed a subject into the playback chamber and allowed him or her to acclimate to this environment for 1 min. After this acclimation period, we played a stimulus approximately every 30 s, and scored whether subjects gave long calls following playback. Responses were scored using the same Hypercard program used to run playbacks. This program allowed us to monitor accurately the time elapsed following the offset of the stimulus and, when a key was pressed, provided a time stamp marking the occurrence of a response. As tamarins sometimes spontaneously produce long calls when isolated, we required tamarins to be silent for 10 s before presenting any stimuli at the start of the experiment and before playing any stimuli. A playback session lasted ca. 5 min.

We played back call sequences consisting of normal CLCs alternating with calls modified in one of the eight ways described above. Stimuli were presented in a nonrandom order so that we could evenly sample responses to normal calls and manipulated calls and space exemplars from the same animal as far apart as possible (Fig. 1).

For each condition, we tested subjects with Call set 1A and then, after waiting at least 2 days, with Call set 1B. Subjects were run on Call sets 2A and 2B (e.g. global reversed condition) 1 week after the completion of Call sets 1A and 1B (e.g. local reversed condition). This test sequence (Call set 1, then Call set 2) was then repeated ca. 2 weeks later. Thus, every subject was tested on each call set (both A and B) twice. Consequently, the following conditions were paired: local reversed with global reversed, low-pitch shift with high-pitch shift, amplitude-modulated noise with gated noise, and intersyllable interval contracted with intersyllable interval expanded. Although a balanced order of testing would have been ideal, the presentation of both normal and manipulated calls within each session provides an adequate control for order effects. Our controls were within session and a new set of calls was used in each condition.

If subjects did not call to at least two of the playbacks (normal or manipulated) in a call set, then they were rerun on that call set at a later date (usually the following day). The rerunning of subjects occurred infrequently. In the 256 experimental sessions reported here (eight subjects participated in 32 sessions each), subjects were rerun only eight times (or ca. 3%). Furthermore, four of these reruns occurred in the first half of the experiments and six of the eight subjects were rerun. Thus, our results are unlikely to have been biased by the small number of subjects that were rerun.

Response Criteria

During a session, subjects produced orienting and scanning behaviours and gave calls (e.g. chirps) other than long calls. However, according to the behavioural observations of Cleveland & Snowdon (1982) and observations from our previous experiments (Ghazanfar et al. 2001), the only unambiguous response to playbacks of CLCs is antiphonal calling with CLCs by the subject. Thus, antiphonal calling was our sole response measure. If a long call was produced just prior to or concurrently with the onset of a stimulus presentation (as judged by the experimenter), it was scored as 'ambiguous' and counted as a 'no response' in our analyses. A long call produced by a subject within 5 s of the offset of a stimulus was scored as an antiphonal call, while long calls produced beyond 5 s poststimulus and until the next stimulus presentation were scored as 'spontaneous'. Antiphonal calls were scored on a yes/no basis. To test the validity of this criterion, we compared the temporal distribution of long calls produced by our subjects after playbacks of CLCs with the distribution of long calls produced following a 2.5-s period of silence. An audio file of 2.5 s of silence was used to mimic the average length of a long call, thereby allowing us to recreate the exact experimental conditions used normally when normal long calls are alternated with manipulated long calls. Playbacks of normal CLCs and periods of silence were arranged in a playback sequence similar to all other conditions. Subjects were tested only once on each version of the call set.

Analysis

The number of antiphonal calls in each condition were entered into a repeated measures ANOVA with session number and call sets as between-subject factors and call type (Normal versus Manipulated) as a within-subject factor. Effect sizes are reported using the f statistic (Rosenthal & Rosnow 1991). Effect sizes indicate the magnitude of the effect independent of sample size by assessing the percentage of overlap between two distributions. The larger the effect size, the more robust is the difference between two means (i.e. the less overlap). According to Cohen (1988), effect sizes of 0.8 and above are large, 0.5–0.7 are medium and 0.2–0.4 are small. As
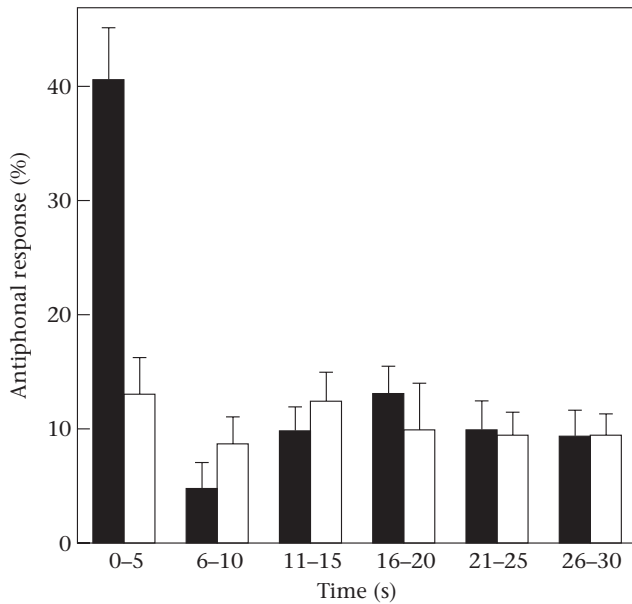
Figure 2. Temporal distribution of the percentage of long calls given by subjects following playbacks of normal long calls (■) or 2.5 s of silence (□). Data are $\bar{X} \pm 1$ SEM.

there were fewer than 10 subjects in all conditions, we also report the results of a nonparametric statistical test, the sign test, which is based on the proportion of subjects antiphonal calling more or less frequently to unmanipulated signals versus manipulated ones. The average number of antiphonal calls to either normal calls or manipulated calls was divided by 10 (the number of possible antiphonal calls in either category) to give a percentage. Data are presented in terms of means ± SE.

RESULTS

Response Criterion

To test the validity of the 5-s criterion, we recorded the rate of long calling in the 0–30 s time period following a playback. The highest rate of calling was in the 0–5 s time bin (Fig. 2). All other time bins had roughly equal rates of long calling and were not distinguishable from the level of spontaneous long calling seen during the silent control period (Fig. 2). Thus, we conclude that our 5-s criterion is a robust measure of antiphonal calling in tamarins.

Local and Global Spectral Contours

By time reversing each syllable of a long call ('local reversal'), we changed the direction of FM sweeps and the amplitude envelopes (Fig. 3) but retained temporal order and overall spectral content of each syllable in the call. The level of calling to local-reversed long calls (33.59 ± 4.17%) was comparable to that produced in response to unmanipulated long calls (32.42 ± 4.00%; $F_{1,7}=0.132$, $P=0.728$, $f=0.134$; sign test: $P=0.637$).

Temporally reversing the entire long call ('global reversal') changed not only the FM structure and ampli-

tude envelope, but also reversed the order of syllables (Fig. 3). Surprisingly, this manipulation did not affect the tamarins' antiphonal calling responses. Tamarins called antiphonally to normal calls on 30.08 ± 2.37% of trials and to global-reversed calls on 31.64 ± 3.17% of trials ($F_{1,7}=0.233$, $P=0.644$, $f=0.182$; sign test: $P=0.855$).

These data suggest that socially isolated tamarins do not rely on spectral contour (at either the local or global levels) or syllable order information to identify and respond to the long calls of their conspecifics.

Pitch-shifted Long Calls

In the pitch-shifted conditions, the fundamental frequency of individual CLCs was either lowered by half or doubled in order to compare the responses to these manipulated calls with normal CLCs (Fig. 4). We found no difference in responses to normal (34.06 ± 2.97%) versus lower-pitched calls (28.13 ± 2.89%; $F_{1,7}=1.60$, $P=0.247$, $f=0.478$; sign test: $P=0.145$) nor between normal (31.25 ± 3.17%) versus higher-pitched calls (30.31 ± 3.63%; $F_{1,7}=0.352$, $P=0.572$, $f=0.222$; sign test: $P=0.363$). Thus, in the context of social isolation, we found no evidence that tamarins are more or less responsive to higher- or lower-pitched calls than to normal CLCs.

Amplitude-modulated and Gated White Noise

In the amplitude-modulated stimulus set we replaced all the natural spectral content with white noise, while retaining the species-typical amplitude envelope (Fig. 5). Tamarin antiphonal calling responses were not significantly different to normal calls (28.44 ± 2.91%) versus amplitude-modulated noise (23.43 ± 3.37%; $F_{1,7}=2.732$, $P=0.142$, $f=0.625$; sign test: $P=0.145$).

In the 'gated noise' condition, we replaced long call syllables with blocks of white noise of equivalent duration, thus eliminating the species-typical amplitude envelope (Fig. 6). Tamarins responded significantly less to the gated noise (17.19 ± 2.59%) than to normal CLCs (30.63 ± 2.91%; $F_{1,7}=13.275$, $P=0.008$, $f=1.375$; sign test: $P=0.004$).

Intersyllable Intervals

By eliminating (Fig. 7) or expanding (Fig. 8) the intersyllable interval of long calls, we tested whether tamarins attend to this temporal feature in their antiphonal long-calling behaviour. When the intersyllable interval was eliminated, tamarins called antiphonally more (30.03 ± 3.40%) than to normal calls (24.06 ± 2.59%). This was significant as measured by the ANOVA ($F_{1,8}=6.914$, $P=0.030$), but not by the sign test ($P=0.09$). However, the effect size was large ($f=0.929$; Cohen 1988) which lends support to the conclusion that tamarins call antiphonally more to contracted long calls. Tamarins did not call antiphonally more or less to calls whose intersyllables were expanded to twice the maximum observed in our colony (25.63 ± 3.51 versus 28.13 ± 3.37%; $F_{1,7}=0.491$, $P=0.506$, $f=0.266$; sign test: $P=0.855$).
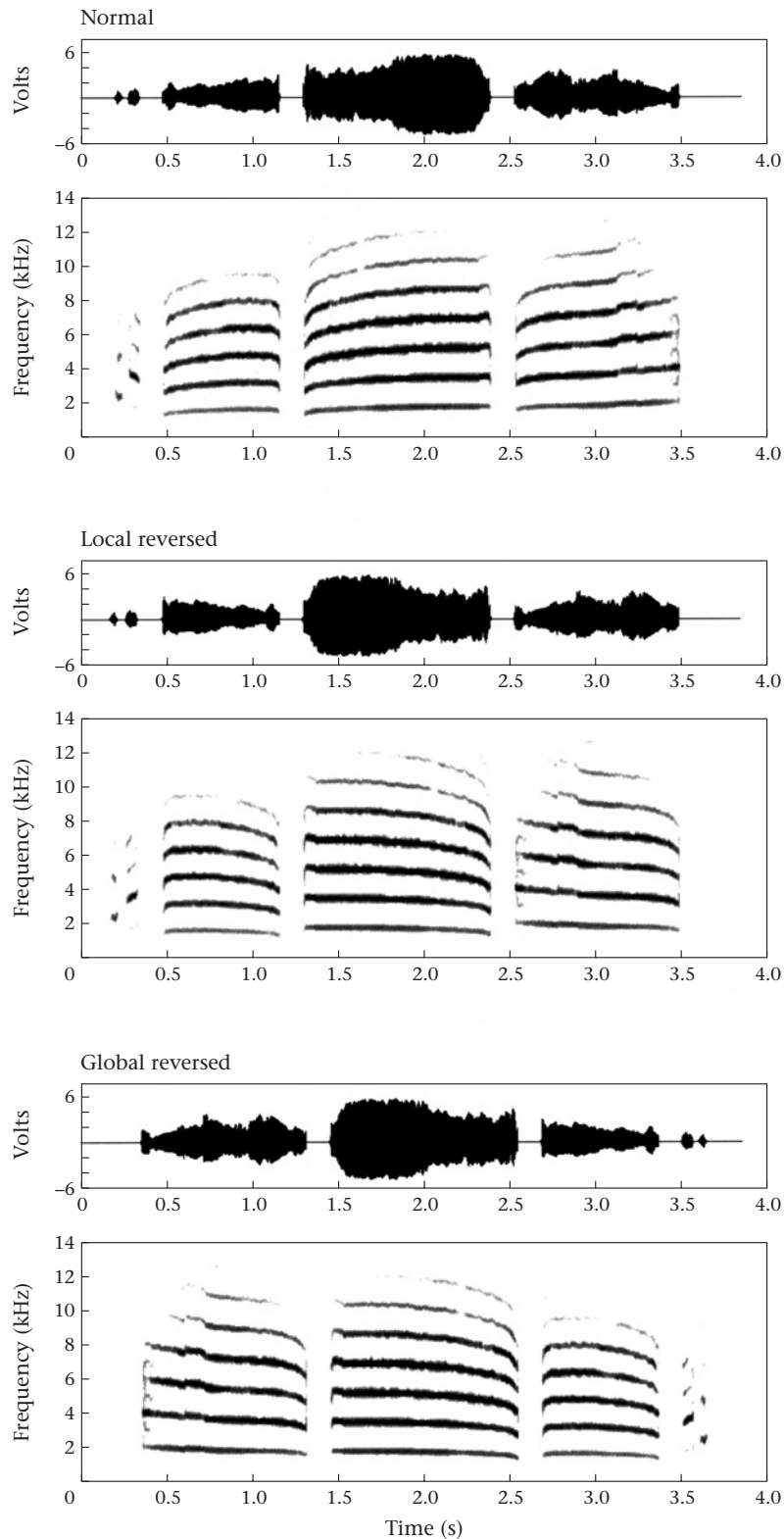
Figure 3. Time–amplitude waveform and spectrogram of a normal, local-reversed and global-reversed long call exemplar (see text for details of manipulations). Spectrograms were generated from a 256-pt Fast Fourier Transform (Hanning window applied).

Subjects often interrupted playbacks of expanded long calls (43% of antiphonal responses) significantly more than they interrupted playbacks of normal calls (11% of

antiphonal responses; one sample t test, $t_{72}=5.596$, $P<0.0001$; all eight individuals, sign test: $P=0.004$; Fig. 9). For expanded long calls, the majority of call interruptions
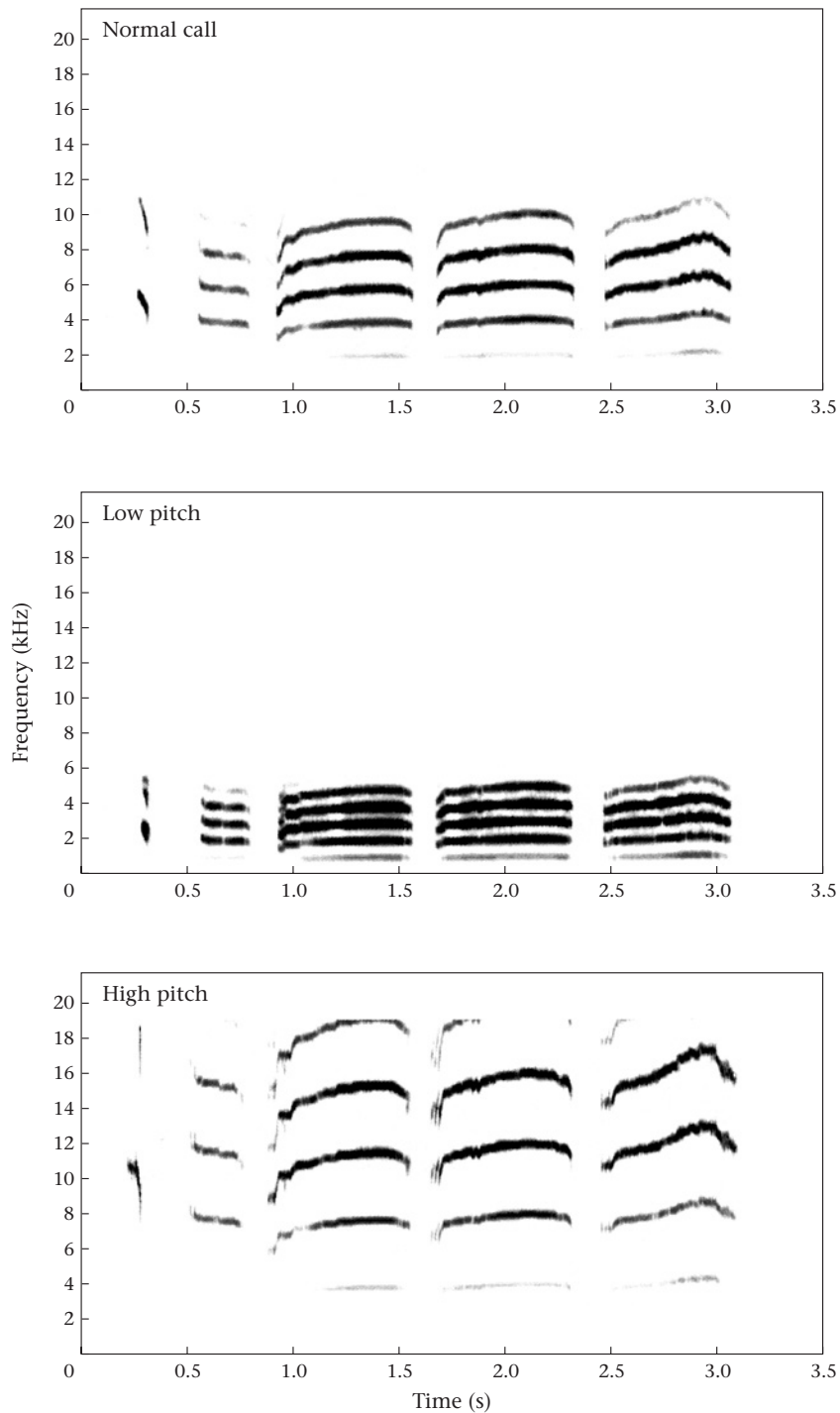
Figure 4. Pitch shift condition. Spectrograms of a normal, low-pitch shifted and high-pitch shifted long call.

occurred right after, or in the middle of, the third syllable (χ^2_3=24.25, P<0.0001; Fig. 9). The same was true for the few call interruptions that occurred during the playbacks of normal CLCs: in the seven interruptions (out of 61 antiphonal responses), five occurred after (or in the middle of) the third syllable and two after (or in the middle of) the fourth syllable.

The results from these intersyllable interval experiments suggest that tamarins may be more likely to call antiphonally to CLCs with higher syllable rates. Increased call duration (syllables and intersyllable intervals) did not seem to be a factor in tamarin antiphonal calling behaviour as stretching the call to almost twice its normal duration did not result in an increase in antiphonal calling rates.
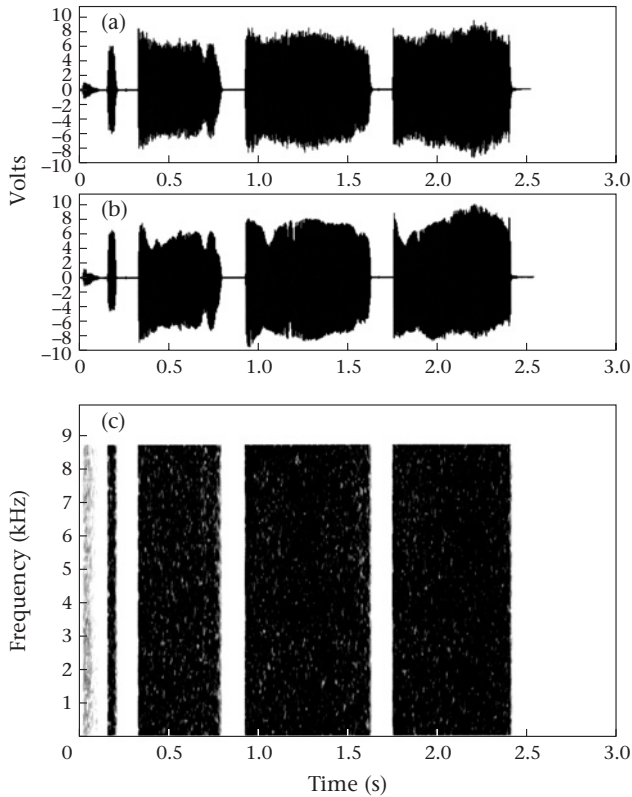
Figure 5. Amplitude-modulated noise condition. (a) Time–amplitude waveforms of an original long call and (b) its synthetic counterpart. (c) Spectrogram of the synthetic call shown above.
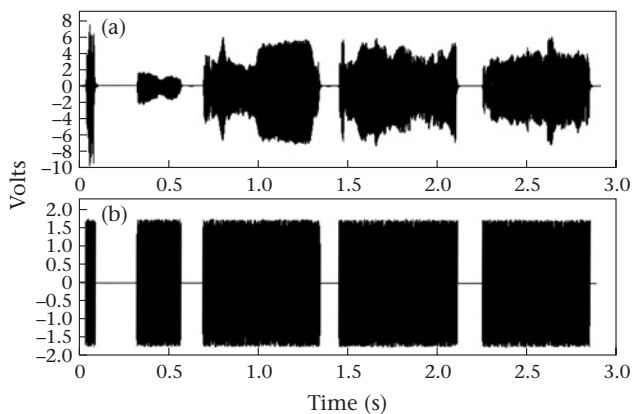

Figure 6. Gated noise condition. Time–amplitude waveforms of (a) an original long call and (b) its synthetic counterpart.

Overall Response Rates

All subjects were exposed to the different experimental conditions in the same order. We controlled for possible order effects by presenting both normal and manipulated calls in each session. A linear regression of the data revealed no significant decline in the responses to normal calls across the eight conditions ($F_{1,262}=0.388$, $P=0.534$). Furthermore, the failure to show a difference in earlier conditions was followed by conditions in which the
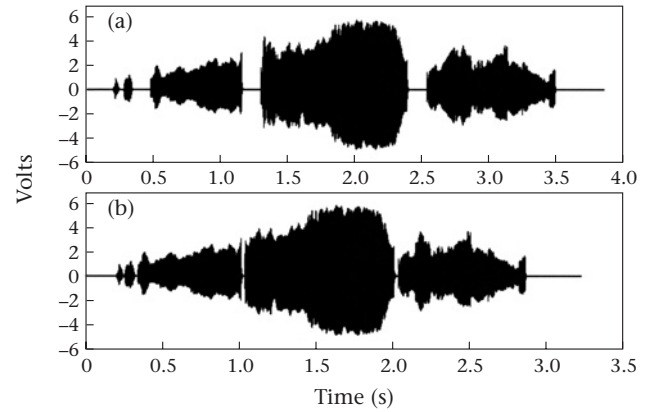

Figure 7. Intersyllable interval contraction condition. (a) Time–amplitude waveforms of an original long call and (b) the same call with the intersyllable interval eliminated.
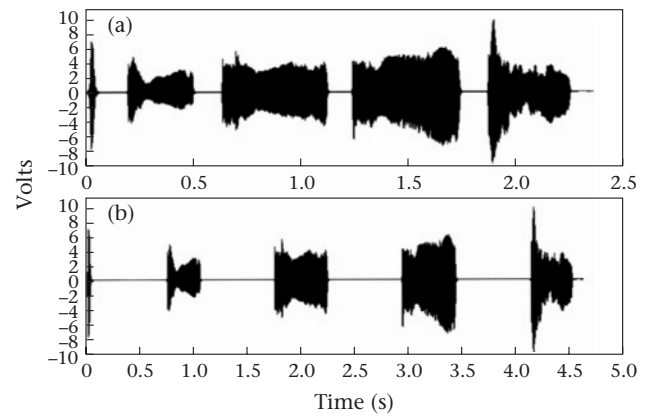

Figure 8. Intersyllable interval expansion condition. (a) Time–amplitude waveforms of an original long call and (b) the same call with the intersyllable interval expanded.

tamarins showed significant differences in antiphonal calling rates relative to normal calls.

DISCUSSION

We investigated which acoustic features of the tamarin combination long call (CLC) are necessary to elicit antiphonal long calls. We found that reversing the entire call, or its individual syllables (but preserving their temporal order) did not affect our subjects' rates of antiphonal calling. Raising or lowering the fundamental frequency and replacing all species-typical spectral content of long calls with white noise also did not affect antiphonal calling rates. However, when we eliminated the species-typical amplitude envelope, we elicited significantly lower rates of antiphonal calling. In addition, increasing the syllable rate (by eliminating intersyllable intervals) increased our subjects' antiphonal calling rates; this effect, while large according to effect size estimation ($f=0.929$; Cohen 1988), was not significant by the sign test. It should, therefore, be treated with caution. Decreasing syllable rate (by expanding the intersyllable interval) did not affect antiphonal calling rates, although tamarins tended to interrupt these calls at greater rates than they
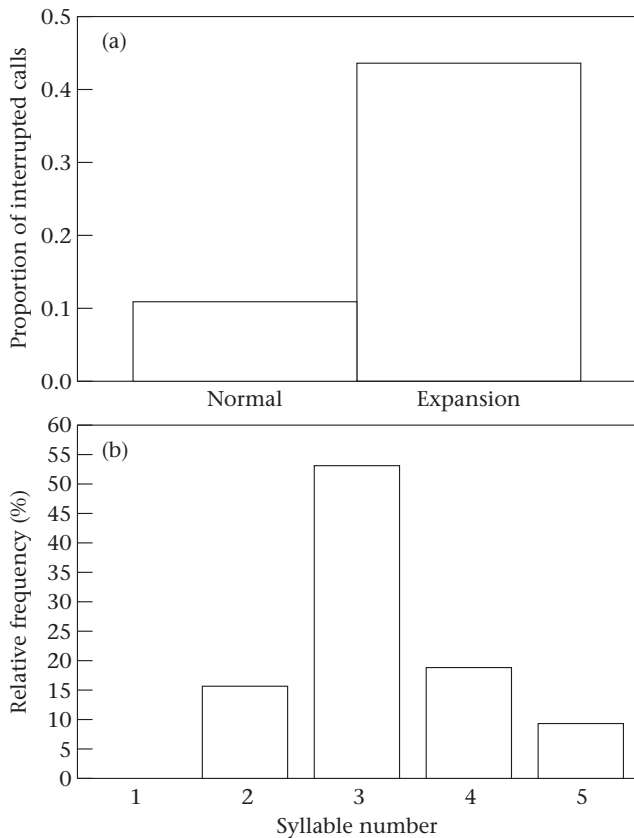
Figure 9. Call interruptions during the intersyllable interval expansion condition. (a) Proportion of calls interrupted with antiphonal calls during playbacks of normal and expanded long calls. (b) Distribution of interruption locations during playbacks of expanded calls.

interrupted normal CLCs. These call interruptions almost always occurred after (or in the middle of) the third syllable of the playback of a five-syllable CLC.

Long Call Recognition

Our previous investigation into the units of perception in tamarin long calls revealed that tamarins called antiphonally more to whole CLCs than to either the chirp or whistle syllables presented in isolation (Ghazanfar et al. 2001). Furthermore, tamarins responded at equivalent rates to artificial sequences of chirps with species-average intersyllable intervals, suggesting that the combination of chirps and whistles is not necessary for eliciting antiphonal calls (Ghazanfar et al. 2001). As chirps have a shorter duration and a higher fundamental frequency than whistles (Cleveland & Snowdon 1982), we can also conclude that the overall amount of acoustic energy per unit time is not a critical feature either. Together with the current findings, we conclude that for socially isolated tamarins temporal cues are largely responsible for eliciting antiphonal calls. Specifically, antiphonal calling is mediated by species-typical amplitude envelopes and, possibly, intersyllable intervals. Furthermore, based on the call interruption data, antiphonal calling appears to

depend on a threshold number of syllables, such that acoustic stimuli with fewer than three syllables elicit significantly lower rates of calling (see also Ghazanfar et al. 2001). An alternative hypothesis would be that the interruption effect we describe is an epiphenomenon of the response latency. Because there is an approximately 5-s response latency following the onset of a long call (Fig. 2), it is possible that the expanded calls are interrupted around the third syllable (which is on average 5 s from the onset of the call) by chance alone. However, in the small number of cases in which the normal, unexpanded call was interrupted, these interruptions always occurred during or after the third syllable, suggesting that syllable number is the important factor. Future experiments in which response latency is measured in each condition could better distinguish between these two hypotheses.

While tamarins do not seem to use spectral cues for antiphonal calling in the context of social isolation, they appear to use this acoustic information in other contexts. Acoustical analyses have shown that individual, gender and group identities are encoded in the spectral features of long calls (Weiss et al. 2001). Playback experiments reveal that individuals are sensitive to harmonic perturbations, including mistuning the harmonics (Weiss & Hauser 2002). Experiments in other taxa have shown that different acoustic properties of communication signals may encode different kinds of biologically significant information (Marler 1961; Gerhardt 1992).

Multiple Acoustic Cues in Vocal Recognition

Organisms commonly use multiple acoustic features in classifying their vocal signals (e.g. Doherty 1985; Nelson 1988). For example, field sparrows, *Spizella pusilla*, are sensitive to changes in five different acoustic features of their songs, but tend to weight features differently (Nelson 1988). Phrase structure and frequency tend to be more important than internote intervals or note duration for song recognition in this species. Similarly, when female crickets listen to male song, their phonotactic behaviour is influenced by an interaction of several temporal features (Doherty 1985).

Our data indicate two possible cues that tamarins may use to identify conspecific long calls. First, tamarins treat amplitude-modulated noise (modelled after their long calls) as equivalent to unmanipulated CLCs, at least with respect to their antiphonal calling response. This strongly suggests that, at least for the context in which we tested our subjects, the temporal envelope is sufficient for recognizing the call as a species-typical CLC, one that leads to an antiphonal response. The fact that a severely degraded signal can elicit an antiphonal response in tamarins is certainly not unique. For example, the blackcap, *Sylvia atricapilla*, gives robust territorial responses to conspecific songs that lack species-typical amplitude envelopes or intersyllable intervals, or have been time reversed (Mathevon & Aubin 2001). Second, tamarins may use the intersyllable interval as a cue: they responded at a higher rate to contracted CLCs than to normal calls,

but did not respond at lower rates for the expanded intersyllable interval condition.

Temporal Cues

Across a diverse range of taxa, the temporal cues of animal vocal signals play an important role in species recognition (Becker 1982; Doherty & Hoy 1985; Gerhardt 1994). In a study of two closely related frog species, *Hyla cinerea* and *H. gratiosa*, Gerhardt (1981) found, using playbacks of normal versus reversed conspecific calls, that the amplitude envelope plays an important role in call recognition for *H. gratiosa* females, but not *H. cinerea* females. In indigo buntings, *Passerina cyanea*, Emlen (1972) found that increasing or decreasing the inter-syllable interval between song notes decreased male responsiveness to conspecific songs.

In the case of nonhuman primates, we know relatively little about the mechanisms underlying species recognition of vocal signals. Virtually nothing is known about the acoustic cues that primates use to assign exemplars into vocal types, and of those studies that have attempted to extract relevant features, practically all focus on spectral parameters. For example, studies have suggested the importance of spectral cues in call discrimination (Owren 1990a, b), individual recognition of conspecifics (Rendall et al. 1996, 1998), call convergence (Sugiura 1998), call development (Gouzoules & Gouzoules 1989) and the relationship between signal design and motivational states (Hauser 1993; Gouzoules & Gouzoules 2000). Thus, we have gained much from studying the spectral cues of primate vocal signals.

The few studies that have examined temporal features in primates suggest that such features may also play an important role in communication. For example, field studies of Japanese macaques, *Macaca fuscata*, revealed two types of coo calls that differed in terms of behavioural context and could be distinguished acoustically by the temporal position of the peak frequency (Green 1975). In captivity, Japanese macaques can use this temporal cue as well as several other cues to classify perceptually their coo vocalizations (Zoloth et al. 1979; May et al. 1988; Le Prell & Moody 1997, 2000). To date, only two playback studies have explicitly explored the acoustic features that primates use to distinguish conspecific from nonconspecific sounds (Hauser et al. 1998; Ghazanfar et al. 2002). Building upon their initial findings using the head-orienting bias as an assay (Hauser & Andersson 1994), Hauser et al. (1998) tested the role of interpulse interval on call recognition in rhesus monkeys, *Macaca mulatta*. Playback experiments using temporally manipulated exemplars of a subset of these calls (shrill barks, grunts and copulation screams) showed that expanding or contracting the interpulse interval beyond the species-typical range eliminated the right-ear orienting bias normally seen for shrill barks and grunts, but not for copulation screams. Sensitivity in rhesus monkeys to the shape of the amplitude envelopes for shrill barks and harmonic arch calls was also observed using time-reversed calls and the head-orienting assay (Ghazanfar et al. 2002). Our present results

show that the amplitude envelope, the number of syllables and, possibly, the intersyllable interval are important cues in long-call recognition by cottontop tamarins.

References

Andrew, R. J. 1963. Origins and evolution of the calls and facial expressions of the primates. *Behaviour*, **20**, 1–108.

Becker, P. 1982. The coding of species-specific characteristics in birds. In: *Acoustic Communication in Birds. Vol. 1* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 213–252. New York: Academic Press.

Boinski, S. 2000. Social manipulation within and between troops mediates primate group movement. In: *On the Move: How and Why Animals Travel in Groups* (Ed. by S. Boinski & P. A. Garber), pp. 421–469. Chicago: University of Chicago Press.

Bosch, J., Rand, A. S. & Ryan, M. J. 2000. Acoustic competition in *Physalaemus pustulosus*, a differential response to calls of relative frequency. *Ethology*, **106**, 865–871.

Cheney, D. L., Seyfarth, R. M. & Palombit, R. 1996. The function and mechanisms underlying baboon contact barks. *Animal Behaviour*, **52**, 507–518.

Cleveland, J. & Snowdon, C. T. 1982. The complex vocal repertoire of the adult cotton-top tamarin, *Saguinus oedipus oedipus*. *Zeitschrift für Tierpsychologie*, **58**, 231–270.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, New Jersey: L. Erlbaum.

Doherty, J. A. 1985. Trade-off phenomena in calling song recognition and phonotaxis in the cricket, *Gryllus bimaculatus* (Orthoptera, Gryllidae). *Journal of Comparative Physiology*, **156**, 787–801.

Doherty, J. A. & Hoy, R. R. 1985. Communication in insects. III. the auditory behavior of crickets: some views of genetic coupling, song recognition, and predator detection. *Quarterly Review of Biology*, **60**, 457–472.

Emlen, S. T. 1972. An experimental analysis of the parameters of bird song eliciting species recognition. *Behaviour*, **41**, 130–171.

Ewing, A. M. 1989. *Arthropod Bioacoustics, Neurobiology and Behavior*. Ithaca, New York: Cornell University Press.

Fischer, J. 1998. Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*, **55**, 799–807.

Fischer, J., Metz, M., Cheney, D. L. & Seyfarth, R. M. 2001. Baboon responses to graded bark variants. *Animal Behaviour*, **61**, 925–931.

Fitch, W. T. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, **102**, 1213–1222.

Fitch, W. T. & Kelley, J. P. 2000. Perception of vocal tract resonances by whooping cranes, *Grus americana*. *Ethology*, **106**, 559–574.

Fritzsch, B., Ryan, M. J., Wilczynski, W., Hetherington, T. E. & Walkowiak, W. 1988. *The Evolution of the Amphibian Nervous System*. New York: J. Wiley.

Gerhardt, H. C. 1981. Mating call recognition in the barking treefrog (*Hyla gratiosa*): responses to synthetic mating calls and comparisons with the green treefrogs (*Hyla cinerea*). *Journal of Comparative Physiology A*, **144**, 17–25.

Gerhardt, H. C. 1991. Female mate choice in treefrogs: static and dynamic acoustic criteria. *Animal Behaviour*, **42**, 615–635.

Gerhardt, H. C. 1992. Multiple messages in acoustic signals. *Seminars in the Neurosciences*, **4**, 391–400.

Gerhardt, H. C. 1994. The evolution of vocalization in frogs and toads. *Annual Review of Ecology and Systematics*, **25**, 293–324.

Ghazanfar, A. A., Flombaum, J. I., Miller, C. T. & Hauser, M. D. 2001. The units of perception in the antiphonal calling behavior of cotton-top tamarins. (*Saguinus oedipus*): playback experiments with long calls. *Journal of Comparative Physiology A*, **187**, 27–35.

Ghazanfar, A. A., Smith-Rohrberg, D. & Hauser, M. D. 2002. The role of temporal cues in conspecific vocal recognition: rhesus monkey orienting asymmetries to reversed calls. *Brain, Behavior, and Evolution*, **58**, 163–172.

Gouzoules, H. & Gouzoules, S. 1989. Design features and developmental modifications of pigtail macaque, *Macaca nemestrina*, agonistic screams. *Animal Behaviour*, **37**, 383–401.

Gouzoules, H. & Gouzoules, S. 2000. Agonistic screams differ among four species of macaques: the significance of motivation-structure rules. *Animal Behaviour*, **59**, 501–512.

Green, S. 1975. Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): a field study. In: *Primate Behavior, Vol. 4* (Ed. by L. A. Rosenblum), pp. 1–102. New York: Academic Press.

Hauser, M. D. 1993. The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight, and social context. *American Naturalist*, **142**, 528–542.

Hauser, M. D. & Andersson, K. 1994. Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. *Proceedings of the National Academy of Sciences, U.S.A.*, **91**, 3946–3948.

Hauser, M. D., Agnetta, B. & Perez, C. 1998. Orienting asymmetries in rhesus monkeys: effect of time-domain changes on acoustic perception. *Animal Behaviour*, **56**, 41–47.

Holland, J., Dabelsteen, T. & Paris, A. L. 2000. Coding in the song of the wren: importance of rhythmicity, syntax and element structure. *Animal Behaviour*, **60**, 463–470.

Kroodsma, D. E. & Miller, E. H. 1996. *Ecology and Evolution of Acoustic Communication in Birds*. Ithaca, New York: Cornell University Press.

Lengagne, T., Aubin, T., Jouventin, P. & Lauga, J. 2000. Perceptual salience of individually distinctive features in the calls of adult king penguins. *Journal of the Acoustical Society of America*, **107**, 508–516.

Le Prell, C. G. & Moody, D. B. 1997. Perceptual salience of acoustic features of Japanese monkey coo calls. *Journal of Comparative Psychology*, **111**, 261–274.

Le Prell, C. G. & Moody, D. B. 2000. Factors influencing the salience of temporal cues in the discrimination of synthetic Japanese monkey (*Macaca fuscata*) coo calls. *Journal of Experimental Psychology: Animal Behavioral Processes*, **26**, 261–273.

McConnell, P. B. & Snowdon, C. T. 1986. Vocal interactions between unfamiliar groups of captive cotton-top tamarins. *Behaviour*, **97**, 273–296.

Marler, P. 1961. The logical analysis of animal communication. *Journal of Theoretical Biology*, **1**, 295–317.

Marler, P. 1968. Aggregation and dispersal: two functions in primate communication. In: *Primates: Studies in Adaptation and Variability* (Ed. by P. C. Jay), pp. 420–438. New York: Holt, Rinehart & Winston.

Mathevon, N. & Aubin, T. 2001. Sound-based species-specific recognition in the black cap *Sylvia atricapilla* shows high tolerance to signal modifications. *Behaviour*, **138**, 511–524.

May, B., Moody, D. B. & Stebbins, W. C. 1988. The significant features of Japanese macaque coo sounds: a psychophysical study. *Animal Behaviour*, **36**, 1432–1444.

Miller, C. T. & Ghazanfar, A. A. 2002. Meaningful acoustic units in nonhuman primate vocal behavior. In: *The Cognitive Animal* (Ed. by C. Allen, M. Bekoff & G. M. Burghardt), pp. 265–273. Cambridge, Massachusetts: MIT Press.

Miller, C. T., Dibble, E. & Hauser, M. D. 2001. Amodal completion of acoustic signals by a nonhuman primate. *Nature Neuroscience*, **8**, 783–784.

Mitani, J. C. 1985. Responses of gibbons (*Hylobates muelleri*) to self, neighbor, and stranger duets. *International Journal of Primatology*, **6**, 193–200.

Mitani, J. C. & Marler, P. 1989. A phonological analysis of male gibbon singing behavior. *Behaviour*, **109**, 20–45.

Mitani, J. & Nishida, T. 1993. Contexts and social correlates of long distance calling by male chimpanzees. *Animal Behaviour*, **45**, 735–746.

Narins, P. M. & Capranica, R. R. 1978. Communicative significance of the two-note call of the treefrog *Eleutherodactylus coqui*. *Journal of Comparative Physiology A*, **127**, 1–9.

Nelson, D. A. 1988. Feature weighting in species song recognition by the field sparrow (*Spizella pusilla*). *Behaviour*, **88**, 145–189.

Newman, J. D., Lieblich, A., Talmage-Riggs, G. & Symmes, D. 1978. Syllable classification and sequencing in twitter calls of squirrel monkeys (*Saimiri sciureus*). *Zeitschrift für Tierpsychologie*, **47**, 77–88.

Neyman, P. F. 1978. Aspects of the ecology and social organization of free-ranging cotton-top tamarins (*Saguinus oedipus*) and the conservation status of the species. In: *The Biology and Conservation of the Callitrichidae* (Ed. by D. G. Kleiman), pp. 39–71. Washington: Smithsonian Institution Press.

Owren, M. J. 1990a. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): I. Natural calls. *Journal of Comparative Psychology*, **104**, 20–28.

Owren, M. J. 1990b. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): II. Synthetic calls. *Journal of Comparative Psychology*, **104**, 29–40.

Rendall, D., Rodman, P. S. & Emond, R. E. 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, **51**, 1007–1015.

Rendall, D., Owren, M. J. & Rodman, P. S. 1998. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *Journal of the Acoustical Society of America*, **103**, 602–614.

Rendall, D., Cheney, D. L. & Seyfarth, R. M. 2000. Proximate factors mediating ''contact'' calls in adult female baboons (*Papio cynocephalus ursinus*) and their infants. *Journal of Comparative Psychology*, **114**, 36–46.

Robinson, J. G. 1979. An analysis of the organization of vocal communication in the titi monkey *Callicebus moloch*. *Zeitschrift für Tierpsychologie*, **49**, 381–405.

Rosenthal, R. & Rosnow, R. L. 1991. *Essentials of Behavioral Research: Methods and Data Analysis*. New York: McGraw-Hill Publishing.

Ryan, M. J. & Rand, A. S. 1990. The sensory basis of sexual selection for complex calls in the Tungara frog, *Physalaemus pustulosus* (sexual selection for sensory exploitation). *Evolution*, **44**, 305–314.

Ryan, M. J. & Rand, A. S. 1993. Species recognition and sexual selection as a unitary problem in animal communication. *Evolution*, **47**, 647–657.

Ryan, M. J. & Wilczynski, W. 1988. Coevolution of sender and receiver: effect on local mate preferences in cricket frogs. *Science*, **240**, 1786–1788.

Searcy, W. A. & Marler, P. 1981. A test for responsiveness to song structure and programming in female sparrows. *Science*, **213**, 926–928.

Snowdon, C. T., Cleveland, J. & French, J. A. 1983. Responses to context- and individual-specific cues in cotton-top tamarin long calls. *Animal Behaviour*, **31**, 92–101.

Sugiura, H. 1998. Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, **55**, 673–687.

Waser, P. M. 1977. Individual recognition, intragroup cohesion, and intergroup spacing: evidence from sound playback to forest monkeys. *Behaviour*, **60**, 28–74.

Waser, P. M. 1982. The evolution of male loud calls among mangabeys and baboons. In: *Primate Communication* (Ed. by C. T. Snowdon, C. H. Brown & M. R. Peterson), pp. 177–143. New York: Cambridge University Press.

Weiss, D. J. & Hauser, M. D. 2002. Perception of harmonics in the combination long call of cottontop tamarins, *Saguinus oedipus. Animal Behaviour*, **64** (2002) U27.

Weiss, D. J., Garibaldi, B. T. & Hauser, M. D. 2001. The production and perception of long calls in cotton-top tamarins (*Saguinus oedipus*): acoustic analyses and playback experiments. *Journal of Comparative Psychology*, **115**, 258–271.

Zoloth, S. R., Petersen, M. R., Beecher, M. D., Green, S., Marler, P., Moody, D. B. & Stebbins, W. 1979. Species-specific perceptual processing of vocal sounds by monkeys. *Science*, **204**, 870–873.