



# Perception of harmonics in the combination long call of cottontop tamarins, *Saguinus oedipus*

DANIEL J. WEISS & MARC D. HAUSER

Department of Psychology, Harvard University

(Received 18 May 2001; initial acceptance 27 August 2001;  
final acceptance 14 February 2002; MS. number: A9069R)

A number of nonhuman primates produce vocalizations with time-varying harmonic structure. Relatively little is known about whether such spectral information plays a role in call type classification. We address this problem by utilizing acoustic analyses and playback experiments on cottontop tamarins' combination long call, a species-typical vocalization with a characteristic harmonic structure. Specifically, we used habituation–discrimination experiments to test whether particular frequency components, as well as the relationship between components, have an effect on the perception and classification of long calls. In Condition 1, we show that tamarins classify natural and synthetic exemplars of the long call as perceptually similar, thereby allowing us to use synthetics to manipulate components of this signal precisely. In subsequent conditions, we tested the perceptual salience and discriminability of long calls in which we deleted (1) the second harmonic, (2) the fundamental frequency, or (3) all frequencies above the fundamental; we also examined the effects of frequency mistuning by shifting the second harmonic by 1000 Hz. Following habituation to unmanipulated long calls, tamarins did not respond (transferred habituation) to long calls with either a missing fundamental frequency or the second harmonic, but responded (discriminated) to long calls with the upper harmonics eliminated or with the second harmonic mistuned. These studies reveal the importance of harmonic structure in tamarin perception, and highlight the advantages of using synthetic signals for understanding how particular acoustic features drive perceptual classification in nonhuman primates.

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

The vocal repertoires of animals consist of call types with different acoustic morphologies and different functions. When an animal responds to a vocalization, it has extracted particular acoustic features and used these to identify the call type and its associated function. The aim of research on acoustic perception is to show how animals classify the calls within their repertoire by uncovering the necessary and sufficient features for call recognition and discrimination (for reviews, see Hauser 1996; Bradbury & Vehrencamp 1998).

A general approach in studies of vocal perception is first to describe the vocal repertoire in terms of both the contexts for call production as well as the acoustic morphology of such calls (e.g. Jürgens 1982; Cheney & Seyfarth 1984; Hauser 1991; Farabaugh et al. 1992; Boughman 1997; Owren et al. 1997). A second step involves playback experiments, designed to reveal how

animals respond to vocalizations in the absence of visual input (e.g. Cheney & Seyfarth 1980, 1982; Kajikawa & Hasegawa 1996; Rendall et al. 1996; Leonard et al. 1997; Hammerschmidt & Fischer 1998; Sayigh et al. 1999). Given the results of these experiments, it is common, especially in studies of insects, anurans and birds, to manipulate the structure of the call played back, using either natural or synthetic exemplars (e.g. indigo buntings, *Passerina cyanea*: Emlen 1972; budgerigars: Park & Dooling 1986; treefrogs: Gerhardt 1991; Gerhardt & Watson 1995; crickets: Farris et al. 1997; zebra finches, *Taeniopygia guttata*: Lohr & Dooling, 1998; song sparrows, *Melospiza melodia*: Searcy et al. 1999; cricket frogs: Burmeister et al. 1999; bullfrogs: Simmons & Bean 2000; king penguins, *Aptenodytes patagonica*: Lengagne et al. 2000; and cranes: Fitch & Kelley 2000).

Studies of nonhuman primates have generally focused on the first two steps of this research approach, with much less work conducted at the level of step three, the manipulation of natural and synthetic exemplars. Two examples of the importance of this final step come from studies of Japanese macaques, *Macaca fuscata* (e.g. Le Prell

Correspondence and present address: D. J. Weiss, Brain and Cognitive Sciences, University of Rochester, Meliora Hall office 494, Rochester, NY 14627, U.S.A. (email: [dweiss@bcs.rochester.edu](mailto:dweiss@bcs.rochester.edu)). M. D. Hauser is at the Department of Psychology, Harvard University, Cambridge, MA 02138, U.S.A.

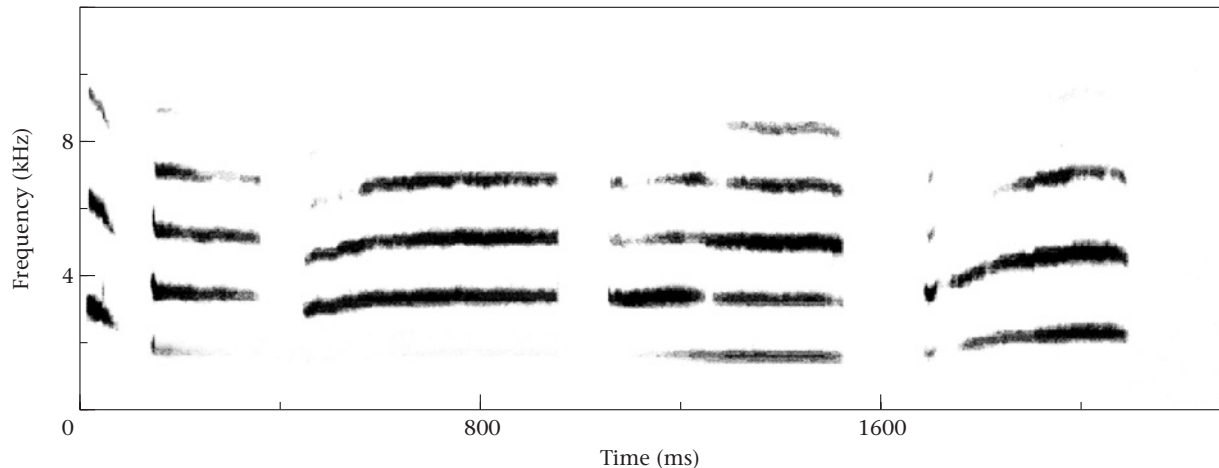


Figure 1. A spectrogram of a five-syllable combination long call.

& Moody 1997, 2000) and vervet monkeys, *Cercopithecus aethiops* (Owren 1990). To illustrate, consider studies conducted by Le Prell & Moody (1997) on the perception of coo calls produced by Japanese macaques. The experimenters created synthetic exemplars of two classes of coo calls (smooth early high, SEH, and smooth late high, SLH) in playback experiments designed to test the perceptual similarity of SEH and SLH coos. Previous studies had shown that the temporal position of the peak frequency was important for coo classification by Japanese macaques (e.g. May et al. 1988). By manipulating the spectral features of the synthetic calls, Le Prell & Moody (1997) found that the relative amplitude of the harmonics was also important for call classification.

In this paper, we focus on the acoustic features underlying call perception in cottontop tamarins. In particular, we explore the perceptual salience of harmonic structure in the classification of the tamarin's combination long call (CLC), a signal that is specially designed for long-distance communication. We select this species and signal for two reasons. First, as illustrated above, our understanding of the mechanisms of perception have been most clearly illuminated by studies involving specialized vocal signals. Second, the cottontop tamarins' vocal repertoire has been studied in great detail, using a variety of acoustic analyses and experimental approaches (Cleveland & Snowdon 1982; Ghazanfar et al. 2001; Weiss et al. 2001). Consequently, we are in an ideal position to begin an in-depth analysis of the acoustic features mediating call perception. To set up the logic of our experiments, we first briefly summarize the results of previous work on the tamarins' CLC.

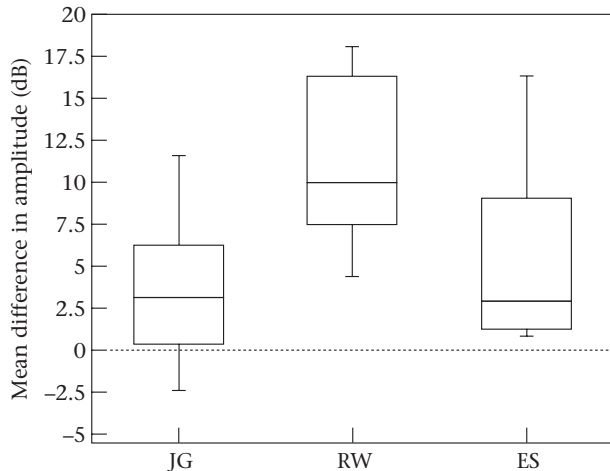
Combination long calls are contact calls that tend to be produced when animals are separated from their group. The call consists of two syllable types, a set of introductory chirps followed by an average of three whistles (Cleveland & Snowdon 1982; for spectrogram, see Fig. 1). Acoustic analyses of the cottontop tamarin CLC indicate that there is potential information about individual, sex and group identity embedded in the structure of the call (Weiss et al. 2001). In addition, habituation–discrimination playback experiments have confirmed

that tamarins can detect changes in caller identity based on acoustic cues alone (Weiss et al. 2001). Phonotaxis experiments have subsequently extended these findings, by demonstrating that tamarins may perceive multiple levels of recognition (including information about cage-mates and familiarity; Miller et al. 2001b). In addition, playback studies have shown that species-typical calls may be most effective in eliciting antiphonal calling (Ghazanfar et al. 2001; Miller et al. 2001a).

The following study begins to explore the contribution of harmonic structure to the perception of the tamarin's CLC. This paper builds on the results of those playback experiments on individual recognition (Weiss et al. 2001). We use an identical playback procedure in which subjects are first habituated to a series of calls from one individual and then presented with a test call. The results from these playbacks are compared with results from the individual recognition conditions.

Our experiments consist of six conditions. The first condition was designed to achieve two aims: (1) to test whether the tamarins would accept our synthetic stimuli as valid, species-typical calls; (2) to assess the perceptual salience of the second harmonic of a CLC by deleting it. Acoustic analyses of five-syllable CLCs (see Garibaldi 1999; Weiss et al. 2001) revealed that the root mean square (RMS) amplitude of the second harmonic was consistently higher than that of the fundamental frequency (see Fig. 2) as well as the higher harmonics (excepting the third harmonic, which also has significant energy). We therefore expected the tamarins to be sensitive to the second harmonic, and respond strongly when it was missing. By ensuring that subjects transfer habituation (i.e. do not respond) to an unmanipulated synthetic call prior to playing a manipulated test call, we can be confident that the tamarins are not responding to acoustic artefacts that may have been introduced in the process of creating the synthetics.

In our previous work, subjects were habituated to a series of calls from one individual and then presented with a novel call from the same individual. Results revealed that subjects consistently (81%) transferred habituation. If our synthetics represent accurate copies of



**Figure 2.** The average amplitude differences between the second harmonic relative to the fundamental frequency across all five syllables of the combination long calls for three subjects (JG, RW and ES).

natural exemplars, then the probability of transferring habituation should be equivalent to our tests using natural exemplars. We test this prediction in Condition 1.

Condition 2 was designed to provide additional evidence that the tamarins accept synthetic calls as valid, species-typical signals. We replicated a condition from a previous experiment (see Weiss et al. 2001) in which subjects were habituated to calls from a male and then played a test call from a female. In that condition, 12 out of 13 subjects responded to the change in caller identity. In Condition 2, we substitute a synthetic version of the test call for the natural exemplar.

Condition 3 was designed to test the salience of the fundamental frequency by deleting it. The fundamental frequency of voiced sounds reflects the rate at which the vocal chords open and close. It is typically the lowest frequency in a complex tone. In humans, the fundamental frequency is known to affect the judgement of both voice quality and recognition (e.g. van Dommelen 1990; Handel 1995). The perception of the missing fundamental (periodicity pitch) is known to have an existence region that extends up to approximately 1400 Hz (Moore 1989). Since the fundamental frequency of the tamarin CLC is close to 2000 Hz, our experiment does not truly address this phenomenon. However, given that the fundamental frequency does play a role in voice recognition in humans (see Handel 1995), it remains an interesting perceptual test for the tamarins. Because there is typically less energy in the fundamental frequency than in the second or third harmonics, one might expect tamarins to treat CLCs with and without the fundamental as perceptually equivalent (see Fig. 2). To provide an additional test that our results were not due to artefacts in the synthetic calls, Condition 4 replicated Condition 3 using natural calls that were hi-pass filtered above the fundamental.

Condition 5 was designed to test whether the deletion of all upper harmonics has an adverse effect on individual recognition. The test stimuli in this condition consisted of natural and synthetic CLCs that were low-pass filtered

above the fundamental frequency, thereby eliminating all of the upper harmonics.

Condition 6 asks whether tamarins perceive a mistuned second harmonic as a meaningful change. There is little known about how primates perceive harmonicity, especially when contrasted with work on humans (e.g. Moore et al. 1986; Hartmann et al. 1990) and songbirds (e.g. Lohr & Dooling 1998). In contrast with work on songbirds and humans, however, our tamarin studies focus not on psychophysically significant sensitivity to harmonic relationships, but rather on meaningful differences that can be detected with our habituation procedures.

The design of our experiments permitted a number of subjects to hear their own calls played back, both in the habituation series and as the test call (including synthetics). Results from a previous study (Weiss et al. 2001) did not find significant differences between responses to own calls versus others' calls. Like the present study, however, the results were limited by a small sample size. Nevertheless, it is worth analysing responses to own calls in comparison with others' calls. It is possible that subjects will perceive own calls as either unfamiliar, familiar but not their own (like a cagemate's call), or as their own. We also explore whether subjects are more responsive to synthetic replicas of their own vocalizations than to those of other individuals.

## METHODS

### Subjects

The subjects consisted of a colony of 15 cottontop tamarins. All colony members except two were born in captivity at the New England Regional Primate Center, Southborough, Massachusetts, U.S.A. and then housed at the Primate Cognitive Neuroscience Laboratory at Harvard University. Subjects were housed in a cage (1.8 × 1.5 × 0.76 m) made of stainless steel wire and Plexiglas; the cage contained tree branches, perches and wooden nestboxes. During the course of these experiments, the colony was housed in six cages. Each cage contained a mated pair; one pair lived with their female offspring and one pair was housed with both a male and female offspring. Their diet consisted of Purina tamarin and marmoset chow, crickets, mealworms, supplemental vitamins and sunflower seeds. This completely balanced diet was supplemented by food received during experiments (typically Noyes banana and nutrasweet pellets, fruit, occasionally Froot Loops<sup>™</sup> and marshmallows). They were fed once a day in the early evening. Subjects had ad libitum access to water.

### Apparatus

The test cage used for the playback experiments was housed in an acoustic chamber (Industrial Acoustics Company, Inc., New York, New York, Model 400-A). During experiments, the tamarins were placed in a wire

and cloth test cage (45 × 45 × 20 cm) with a wire floor. A thin black cotton sheet hung behind the cage. Behind the sheet, an Alesis Monitor One speaker (frequency range, 45–18 000 Hz ± 3 dB) was mounted on a shelf above the box, either directly behind, or to the left or right of centre (the location was changed between experiments to prevent habituation to sounds broadcast from one location).

A video camera (Videolabs Flexcam) was used to record and monitor the sessions. An Alesis RA-100 amplifier drove the speaker. The experimenters watched the session on a monitor outside of the acoustic chamber. The experiment was run and logged using a Hypercard program on a Power Macintosh 7100/80 AV. Calls were played using an Audiomeia II sound card (sampled at 48 kHz) outputting to the Alesis Monitor One speaker. Vocal responses during trials were recorded using a Sennheiser MKH60P48 directional microphone (frequency response 50–20 000 Hz).

### Stimuli

Natural calls used in this experiment were recorded during sessions of another experiment. Additional calls were recorded while the animals were in the playback chamber alone. In both cases, the calls were recorded on a Tascam DAT using a Sennheiser ME-66 microphone (frequency response 40–20 000 Hz).

The calls were hi-pass filtered using Sound Designer II software. To determine the frequency of the filter, we used a spectrogram (1024 Fast Fourier Transform, FFT) to calculate the lowest frequency of the call. After filtering the call at that frequency, we normalized all of the calls to 100% for peak amplitude using Sound Edit 16 v2.0.

Only high-quality calls were selected for use in this experiment. We screened each call by examining the spectrograms (1024 FFT) as well as listening to each exemplar to ensure that they were free of any artefacts (e.g. cage noise, clipping, etc.). In this series of playback experiments, we only used CLCs with five syllables, two 'chirps' and three 'whistles' (see Fig. 1 for a spectrogram of a five-syllable CLC).

### Synthetic Calls

Synthetic calls were created using the SIGNAL digital sound processing and synthesis software (Beeman 1996). The general procedure entailed digitizing (48 kHz sample rate) a natural exemplar and then creating a spectrogram. Next, the fundamental frequency, as well as the frequency and amplitude of each harmonic, were extracted and stored in separate buffers. To manipulate the acoustic parameters of a call, we deleted or shifted (depending on condition) the buffer files for the amplitude and frequency parameters. The beginning and end of each syllable was tapered (using the 'fade out' command on Sound Edit 16) in order to eliminate any artefacts from the synthesis process. We tapered as little as possible for each exemplar, judging by ear whether the artefacts had been removed. Next, the final synthetic exemplar was normalized for peak amplitude to match the level of the

natural calls. Lastly, we played the synthetic signal from a loudspeaker in the acoustic chamber and re-recorded it onto a digital audiotape. This recording was then analysed to ensure that there were no spectral distortions that could introduce energy into the manipulated frequency ranges.

### Call Measurement

Most of the calls in this experiment were drawn from a set of calls that were included in our acoustic analyses of the tamarin CLCs (see Weiss et al. 2001). For this series of experiments, however, we also performed a set of sliding cross-correlations comparing the spectrograms from the synthetic test call against the natural exemplar from which it was created. The cross-correlation analysis was carried out in SIGNAL (Beeman 1996). The process for performing the cross-correlation involved an initial adjustment of all time buffers so that they were equal. Next, digital spectrograms were created from each sound file and then cross-correlated. For each synthetic test stimulus we report the maximum value of the correlation function.

### Procedure

The procedure for our playback experiments was the same as in our previous work with this species (Weiss et al. 2001). We placed a subject into the playback chamber and allowed it to acclimate for 1 min before any recorded sounds were played. After 1 min elapsed, we waited for the subject to face away from the speaker before playing the first call of the habituation set. Habituation stimuli were always randomly presented. If more than one cycle of the habituation set was required during a session, the second presentation of stimuli was presented in a different, randomized order than the first.

For each of the experiments presented in this report, we played a series of calls until the subject failed to respond on three consecutive trials. After three no-response trials, we played the test call, followed by a post-test if necessary (see below). Responses were measured using the following parameters: head turning towards speaker, body orienting towards speaker, movement towards speaker, freezing response (if subject was moving and then froze when the recorded call was played), and vocal responses (producing vocalizations within 5 s of hearing the recorded call). A 'no response' was recorded in the absence of these responses. During sessions, trials with ambiguous responses (as determined by the experimenters while running a session) were treated functionally as 'yes' responses. This conservative approach was adopted in order to minimize the number of trials that needed to be rerun (and thereby minimize exposure to the test stimuli). Intertrial interval was set at a minimum of 10 s and a maximum of 30 s. The session was aborted if we were unable to play back a call within this 30-s window. However, this did not occur during our experiments.

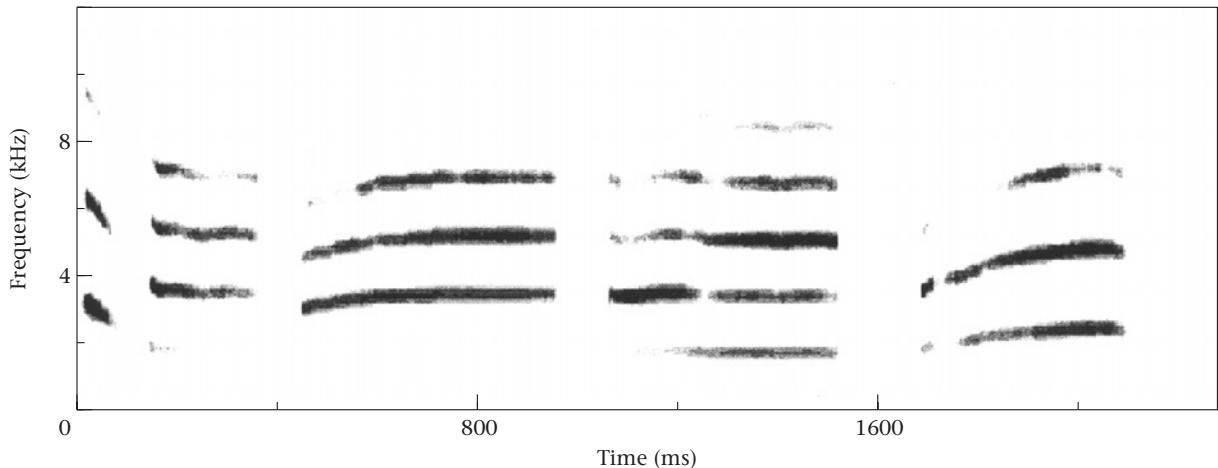


Figure 3. A synthetic replica of a five-syllable combination long call.

Trials started once the subject's face was visible, and turned away from the speaker. In addition, if the subject spontaneously produced a long call we waited at least 5 s before playing back a call.

When subjects failed to respond to the test call, we played back a post-test stimulus. The post-test stimulus, a tamarin scream, represents a call that is both acoustically and functionally different from the habituation and test calls. The scream exemplars were recorded while the medical staff was catching the monkeys for administration of TB shots; before using in playbacks, we normalized each exemplar for amplitude (using the same technique described above). The duration of the post-test screams was within the range of the habituation and test stimuli. The purpose of the post-test was to ensure that the subject had not habituated to the playback set-up in general. Failure to respond to the test stimulus could indicate that the subject had either habituated to the playback set-up in general or had perceptually clustered the habituation and test stimuli into one category. If the subject had habituated to the playback set-up, then it should ignore the post-test stimulus. In contrast, if the failure to respond to the test call was due to the subject's perception of similarity between the habituation series and the test stimulus, then it should respond to the post-test stimulus. Sessions in which subjects failed to respond to the post-test were rerun at a later date.

## Analyses

Because responses to playbacks can be difficult to score in real time, we digitized (Adobe Premiere version 4.2) and scored the last six or seven habituation trials (depending on condition, these included the three 'no response' habituation trials, the test, and the post-test). This ensured that there would be a distribution of both 'yes' and 'no' response trials. These trials were assigned code filenames and then randomized to ensure that the scorers were blind to the condition; once scores were obtained, the master list was consulted to determine the condition. Each trial was marked for the onset and offset

of the playback call as well as the onset and offset of any antiphonal response. The volume was muted during scoring, thereby eliminating any information about trial type. The scorers analysed each of the trials and recorded a 'yes', 'no', or 'ambiguous' (which included bad trials) response; the type of response (e.g. head turn, counter-call, etc.) was recorded as well. Sessions were discarded if scorers assessed any of the three habituation trials (the three trials preceding the test trial) as a 'yes' response. Likewise, sessions in which there was disagreement between scorers on either the three 'no response' habituation trials, the test or post-test stimuli were also discarded. Finally, sessions in which the subject did not respond to both the test and post-test were also discarded. This occurred in less than 7% of all trials.

As mentioned above, the results from these experiments were compared with data obtained from previous playback experiments testing for individual recognition (Weiss et al. 2001). The probability of responding at chance level was calculated from control conditions in which subjects were habituated to calls from one individual and then played back a novel call from that same individual. This represents a conservative measure for chance responding since a novel call was used as the test. We also compared results obtained with manipulated calls with the response rates from test conditions in the aforementioned experiments. In these conditions, subjects were habituated to calls produced by one individual and then tested with a call produced by a novel individual. This allowed us to assess whether the responses observed in this experiment were quantitatively similar to those observed when tamarins detected a change in caller identity.

### *Condition 1: synthetic replica and deleted second harmonic*

Subjects were habituated to a call set consisting of eight natural (unmanipulated) exemplars produced by one individual. Following habituation, subjects were presented with a synthetic replica of one of the calls from the habituation series, selected at random (see Fig. 3 for a spectrogram of a synthetic call). If the subject responded

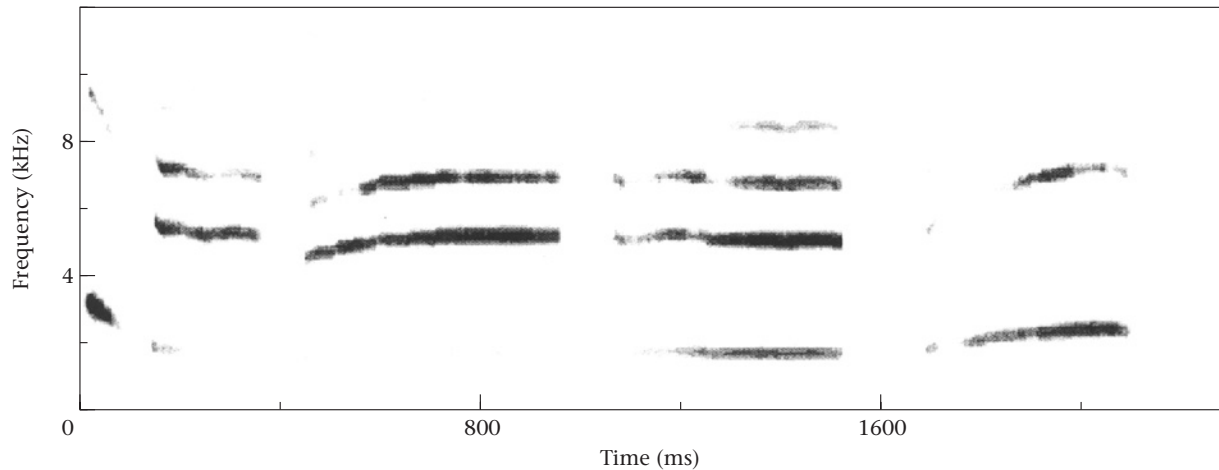


Figure 4. A synthetic replica of a five-syllable combination long call with the second harmonic removed.

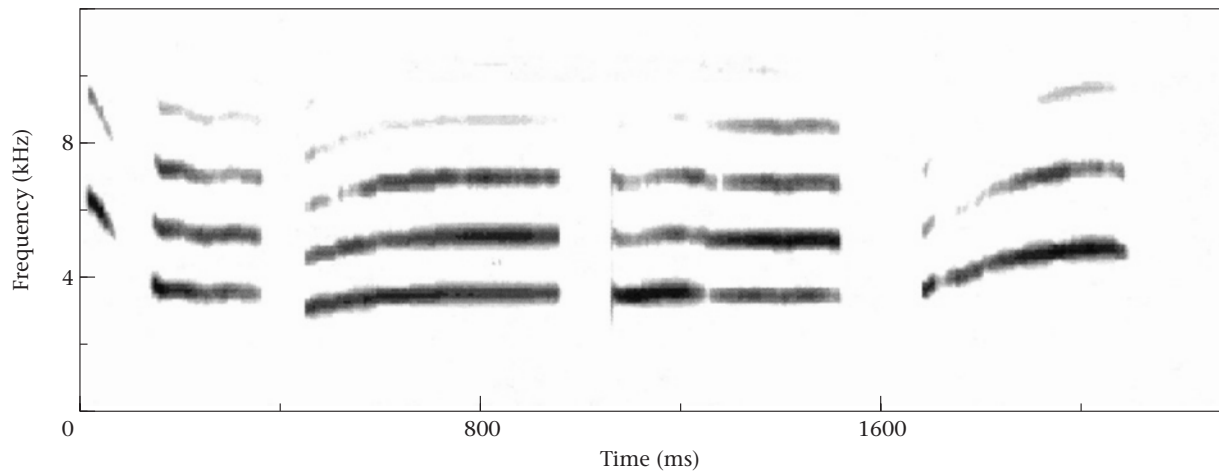


Figure 5. A synthetic replica of a five-syllable combination long call with the fundamental frequency removed.

to the synthetic, the session then ended. If the subject did not respond to the synthetic, we played a synthetic call with a missing second harmonic (see Fig. 4). If the subject responded to this second test trial, the session then ended. If the subject did not respond to the second test call, a post-test stimulus was played. Since this was the first test of the synthetic calls, we repeated this procedure for three different call sets from three different individuals (male RW and females ES and JG).

#### Condition 2: change of identity control

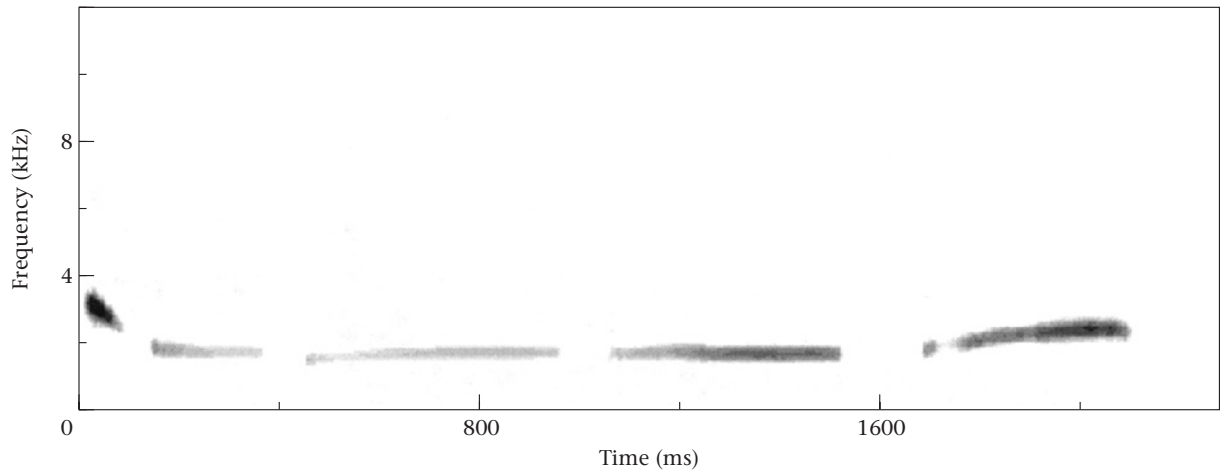
Subjects were habituated to a series of calls from female ES. Subjects were then played a synthetic version of a call from female JG as a test.

#### Condition 3: synthetic calls with no fundamental frequency

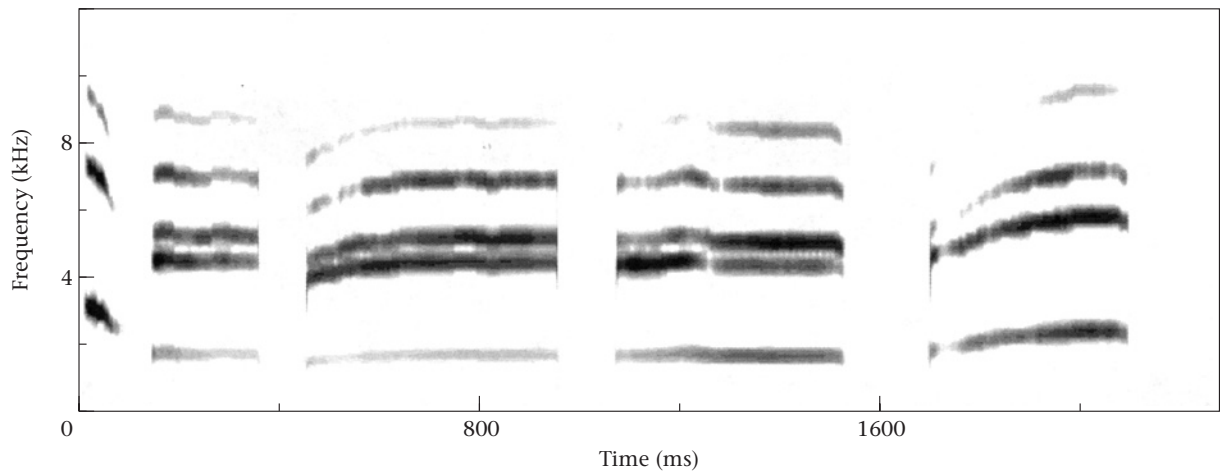
The protocol for this condition was identical to that used in Condition 1 except that we played back a synthetic call consisting of a missing fundamental (see Fig. 5). We used two call sets, one from male RW, and one from female ES.

#### Condition 4: natural calls with no fundamental frequency

The protocol for this condition was similar to that used in Conditions 1 and 3. Following habituation, however, we presented subjects with an additional natural call exemplar (in lieu of the synthetic replica played back in Conditions 1 and 3) from the habituation series. If the subject did not respond to this call, then the same call was played with the fundamental frequency deleted. Calls were edited by first creating spectrograms in Canary (v. 1.2), then hi-pass filtering below the second harmonic using Sound Designer II and, finally, normalizing the amplitude. One difference between this condition and the others is that trials in which subjects responded to the first test call were rerun rather than discarded. This change was designed to maximize data collection since, given the results from previous conditions, we were no longer concerned about the detection of artefacts from our synthetics. We used two call sets, one from male RW and one from female ES. For the ES condition, only 11 subjects were run because of a birth in one of our cage groups.



**Figure 6.** A synthetic replica of a five-syllable combination long call with all of the energy above the fundamental frequency removed.



**Figure 7.** A synthetic replica of a five-syllable combination long call with the second harmonic pitch-shifted by 1000 Hz.

*Condition 5: synthetic and natural exemplars with no harmonics*

In this condition, we made a few changes to our overall experimental procedures in order to facilitate further comparison between synthetic and natural calls. Once subjects habituated to a series of calls from one individual, they were played a test call that consisted of either a natural or synthetic CLC with all energy above the fundamental removed (see Fig. 6). The subjects were randomly divided into two groups, such that half received the natural version first while the other half received the synthetic version first. If subjects did not respond to the first test then they were played the second version as the second test call. If subjects failed to respond to either test, then we played the post-test stimulus. We used two call sets, one from male RW and one from female ES.

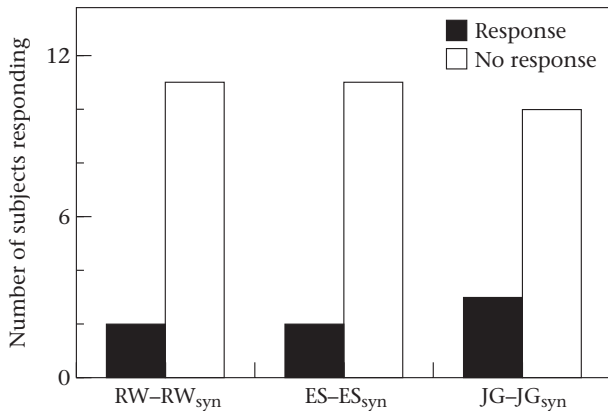
*Condition 6: mistuning the second harmonic*

The protocol for this condition was similar to that used in Condition 1. The test stimulus for this condition

consisted of a synthetic call with a pitch-shifted second harmonic (see Fig. 7). We used two call sets, one from male RW and one from female ES.

We created a mistuned synthetic stimulus in SIGNAL by pitch-shifting the second harmonic by 1000 Hz. We first traced the fundamental frequency from the natural exemplar. Next, we extracted the frequency and amplitude information from the harmonics. Before adding this information to the synthetic fundamental, the frequency buffer for the second harmonic was changed by adding 1000 Hz, thus yielding the mistuned synthetic exemplar. This moved the second harmonic to within 1000 Hz of the third harmonic, as opposed to the average separation of 2000 Hz.

In contrast with previous conditions involving synthetic calls, we reran all trials in which subjects responded to the unmanipulated synthetic stimulus. The reason for this change in methods was to maximize data on the pitch-shifted calls. Given the results from previous conditions, we were confident that the synthetics were perceived as natural calls, and thus discrimination was unlikely to represent artefacts introduced by the process of sound synthesis.



**Figure 8.** Results from Condition 1 showing that most subjects transferred habituation to the synthetic exemplar following habituation to a natural call set.

## RESULTS

### Condition 1: Synthetic Replica and Deleted Second Harmonic

The maximum value of the cross-correlation function comparing the unmanipulated synthetic exemplar against the natural exemplars yielded values of 0.99 for the RW synthetic, 0.98 for the ES synthetic and 0.98 for the JG synthetic. On an acoustic level, therefore, our sound synthesis method worked well. The maximum value of the cross-correlation function comparing the synthetic call with no second harmonic and the natural exemplar was 0.55 for the RW synthetic, 0.70 for the ES synthetic and 0.86 for the JG synthetic.

The average number of trials to habituation was 19.7 (SD=13.9, min=4, max=45) for the RW series, 28.7 calls (SD=27, min=4, max=102) for ES and 18.6 (SD=12.4, min=6, max=54) for JG. There were no significant differences between conditions in the number of trials to habituation (repeated measures analysis of variance, ANOVA:  $F_{2,12}=1.40$ ,  $P<0.27$ ).

Results from the unmanipulated synthetic test calls were as follows (see Fig. 8). For the RW call set, only two out of 13 subjects responded to the test call. These results were not significantly different when compared to control conditions in which subjects were habituated to calls produced by one individual and then tested with a novel call produced by the same individual (see Introduction to Condition 1; binomial test, test proportion: 0.19,  $P<0.54$ ). Likewise, in the ES call set, only two out of 13 subjects responded to the test call (binomial test, test proportion: 0.19,  $P<0.54$ ). For the JG call set, three out of 13 subjects responded to the test call (binomial test, test proportion: 0.19,  $P<0.47$ ). It should be noted that there was no single individual that responded to all three synthetic replicas. There was only one subject that responded to two synthetic replicas. Thus, subjects consistently transferred habituation from the natural exemplars to the synthetic.

Results from the synthetic test calls with the second harmonic removed were as follows. For the RW call set, four out of 11 subjects responded to the test call. This differed significantly from the proportion of subjects that

detected a change of caller identity in same-sex callers (see Weiss et al. 2001; binomial test, test proportion: 0.85,  $P<0.001$ ). For the ES call set, four out of 11 subjects responded to the test call (sign test:  $P<0.55$ ; binomial test, test proportion: 0.85,  $P<0.001$ ). For the JG call set, four out of 10 subjects responded to the test call (sign test:  $P<0.76$ , binomial test, test proportion: 0.85,  $P<0.002$ ).

One of the two subjects that responded to the RW synthetic produced an antiphonal call. Two of the three responding subjects produced antiphonal calls to the JG synthetic, while no subjects called antiphonally to the ES synthetic. The results for antiphonal calling to the test calls with no second harmonic were as follows: two of the four respondents for the RW condition, three out of the four respondents for the ES condition, and two out of four for the JG condition.

### Condition 2: Change of Identity Control

The average number of trials to habituation was 15.7 (SD=9.8, min=5, max=39). This was not significantly different from the number of trials to habituation for the RW set reported in Condition 1 (paired  $t$  test:  $t_{11}=0.55$ ,  $P<0.60$ ). Note that one monkey appeared in this condition that did not participate in Condition 1 and was therefore excluded from this analysis.

In Condition 2, 10 out of 13 subjects responded to the synthetic test call. This result differed significantly from the control condition in which subjects were habituated to calls from male RW and then played back a test call consisting of a novel exemplar from RW (binomial test, test proportion: 0.23,  $P<0.001$ ). This result was not, however, significantly different from the test condition in which subjects were habituated to calls from RW and then played a test call from female ES (binomial test, test proportion: 0.92,  $P<0.08$ ).

Of the 10 respondents in this condition, eight produced antiphonal calls to the synthetic call. This was not significantly different from the rate of antiphonal calling observed in the RW<sub>hab</sub>-ES<sub>test</sub> condition (binomial test, test proportion: 0.46,  $P<0.20$ ).

### Condition 3: Synthetic Calls with No Fundamental Frequency

The maximum value of the cross-correlation function comparing the synthetic call with no fundamental frequency and the natural exemplar was 0.71 for RW's synthetic and 0.86 for ES's synthetic. Both of these values were higher than the values obtained by comparing the synthetic calls with no second harmonic and the natural exemplars.

The average number of calls played back until habituation was 19.9 (SD=14.2, min=5, max=59) for the RW series and 14.8 (SD=19, min=4, max=75) for the ES call set. There were no significant differences between conditions in the number of trials to habituation (repeated measures ANOVA:  $F_{1,12}=0.55$ ,  $P<0.48$ ).

The results from the unmanipulated synthetic test calls were as follows. For the RW call set, only one out of 14

subjects responded to the test call. These results were not significantly different from the control conditions (see Condition 1; binomial test, test proportion: 0.19,  $P < 0.23$ ). In the ES call set, three out of 13 subjects responded to the test call, which was not significantly different from the control condition (binomial test, test proportion: 0.19,  $P < 0.47$ ).

The results from the synthetic test calls with no fundamental frequency were as follows. For the RW call set, six out of 13 subjects responded to the test call; this represents a significantly different level of responding from the proportion of subjects who discriminated between same-sex callers (see Weiss et al. 2001; binomial test, test proportion: 0.85,  $P < 0.002$ ). For the ES call set, three out of 10 subjects responded to the test call; this was also significantly different from the proportion of subjects who discriminated between same-sex callers (binomial test, test proportion: 0.85,  $P < 0.001$ ).

The subject that responded to the RW synthetic did not produce an antiphonal call. In the ES condition, two of the three responding subjects called antiphonally to the test. Results for antiphonal calling to the test calls with no fundamental frequency revealed that two of the six responded in the RW condition, while two out of the three responded in the ES condition.

#### Condition 4: Natural Calls with No Fundamental Frequency

The average number of trials to habituation was 19 (SD=11.7, min=5, max=35) for the RW series and 16.8 (SD=12.3, min=5, max=38) for the ES series. There were no significant differences between conditions in the number of trials to habituation (repeated measures ANOVA,  $F_{1,10}=0.06$ ,  $P < 0.81$ ).

For the RW call set, four out of 14 subjects responded to a test call with no fundamental frequency. This was not significantly different from the results using the synthetic call with no fundamental frequency (paired sign test:  $P < 0.51$ ). All four respondents in this condition also responded to the RW synthetic call with no fundamental. For the ES call set, four out of 11 subjects responded to the test call. These results are similar to the results from the ES synthetic call with no fundamental (paired sign test:  $P < 1.0$ ). Of the four respondents, two responded to the synthetic version as well. However, there were no respondents that responded to both the RW and ES calls with no fundamental.

In the RW test, one of the four respondents called antiphonally, while three out of four respondents called antiphonally in the ES condition.

#### Condition 5: Synthetic and Natural Exemplars with No Harmonics

The maximum value of the cross-correlation function comparing the synthetic call with the fundamental frequency alone and the natural exemplar from which it was created was 0.48 for the RW synthetic and 0.39 for the ES synthetic.

The average number of trials to habituation was 21 (SD=17.9, min=4, max=55) for the RW series and 21.3 (SD=13.7, min=4, max=42) for the ES series. There were no significant differences between conditions in the number of trials to habituation (repeated measures ANOVA:  $F_{1,12}=0.02$ ,  $P < 0.97$ ).

For the RW call set, six out of 13 subjects responded to the first test call. Three out of seven subjects responded to the synthetic stimulus, while three out of six subjects responded to the natural stimulus. For the second test (given only to subjects that did not respond to the first test), four out of seven subjects responded (two to synthetic calls and two to real calls). Thus, overall, 10 out of 13 subjects responded to calls with no energy above the fundamental. This result did not differ significantly from the proportion of subjects responding in the same-sex change of identity conditions (see Weiss et al. 2001; binomial test, test proportion: 0.85,  $P < 0.31$ ).

For the ES call set, 10 out of 13 subjects responded to the first test call. Five out of seven subjects responded to the synthetic, while five out of six subjects responded to the natural call. For the second test, none of the three subjects responded. Thus, the overall number of subjects responding was 10 out of 13, which was identical to the RW condition and thus not significantly different from the response to same-sex change of identity conditions (binomial test, test proportion: 0.85,  $P < 0.31$ ).

In the RW series, five out of 10 respondents called antiphonally to the test call. In the ES series, seven out of 10 respondents called antiphonally to the test call.

#### Condition 6: Mistuning the Second Harmonic

The maximum value of the cross-correlation function comparing the mistuned synthetic call and the natural exemplar from which it was created was 0.46 for the RW synthetic and 0.50 for the ES synthetic.

The average number of trials to habituation was 14 (SD=9.4, min=4, max=30) for the RW series and 18.5 (SD=22.6, min=4, max=90) for the ES series. These results were not significantly different (paired  $t$  test:  $t_{12}=0.69$ ,  $P < 0.51$ ).

For the RW series, 11 out of 14 subjects responded to the mistuned synthetic test call. This did not differ significantly from the proportion of subjects that detected a change of caller identity in same-sex callers (see Weiss et al. 2001; binomial test, test proportion: 0.85,  $P < 0.36$ ). For the ES series, nine out of 13 subjects responded to the mistuned synthetic test call, which was also not significantly different from the aforementioned condition (binomial test, test proportion: 0.85,  $P < 0.12$ ).

For the RW condition, six out of 11 respondents called antiphonally to the manipulated test call, while in the ES condition, three out of nine respondents called back.

#### Playbacks of Own Calls

Across all conditions, subjects who heard their own calls during the habituation series took an average of 15.8 trials (SD=8.5) to habituate. This did not differ significantly from the number of trials those same individuals

required in order to habituate to call sets from other individuals ( $\bar{X} \pm SD = 14.1 \pm 10.4$ ; paired  $t$  test:  $t_9 = 0.43$ ,  $P < 0.68$ ).

Subjects responded to two out of eight trials in which they heard their own, unmanipulated synthetic call. This result was not statistically significant when compared to the proportion of responses by these same subjects to other individual's synthetic calls (binomial test, test proportion: 0.50,  $P < 0.29$ ).

For subjects who heard their own manipulated call (natural or synthetic), five out of nine subjects responded. This was not significantly different from the proportion of remaining subjects responding in those conditions (binomial test, test proportion: 0.45,  $P < 0.38$ ). Similarly, this pattern of response did not differ statistically from the proportion of responses by the same subjects to other individual's manipulated calls (binomial test, test proportion: 0.33,  $P < 0.14$ ).

## DISCUSSION

The main findings from these experiments are that cottontop tamarins classify natural and synthetic exemplars of the CLC as perceptually equivalent, and when assessing the identity of the caller, rely on harmonic structure and harmonic relations, but not on the presence of a single harmonic frequency. Following habituation to a series of long calls from one individual, the tamarins did not significantly respond to manipulated synthetic or natural calls with the fundamental frequency or second harmonic deleted. However, the tamarins did significantly respond to synthetic test calls involving changes in caller identity, deletion of all harmonics above the fundamental, and mistuning of the second harmonic by 1000 Hz. The latter finding suggests that tamarins are perceptually sensitive to the relationship between harmonics. These results are consistent with the earlier reports on Japanese macaques (May et al. 1989) in which removing the harmonics of a coo above the fundamental had a greater effect than removing the fundamental alone. However, our results contrast with more recent findings suggesting that relative harmonic amplitudes play a role in coo perception (Le Prell & Moody 1997). We elaborate on some of our findings below.

One surprising result was that the tamarins treated unmanipulated CLCs as perceptually equivalent to CLCs with deleted fundamentals or second harmonics. However, lack of response to these manipulations might be due to our experimental assay. It is likely that subjects would show evidence of discrimination if the same contrasts were explored using a psychophysical paradigm. Our results suggest that when tamarins classify CLCs by caller identity, they need not rely on the fundamental frequency and second harmonic as necessary acoustic cues. Thus, the tamarins' CLC is likely to carry a suite of acoustic features available for individual recognition (see Weiss et al. 2001), rendering the absence of one harmonic less important in assessing caller identity. Redundancy of information is, apparently, a robust property of many animal communication systems (see Emlen 1972; Park

& Dooling 1986; Gerhardt 1991, 1992; Bradbury & Vehrencamp 1998).

Research has shown that humans can discriminate harmonically structured signals with a single mistuned frequency peak (e.g. Moore et al. 1986; Hartmann et al. 1990). Similarly, zebra finches and budgerigars can detect mistuned harmonics, and with greater acuity than humans (Lohr & Dooling 1998). Our experiment does not address the issue of acuity. Rather, we focused on whether larger perturbations of harmonic structure represent a meaningful difference for tamarins in perceiving long calls. Experiments with bullfrogs (Simmons & Bean 2000) reveal that males modify their response to synthetic calls that have one mistuned harmonic. Those authors claim that the ability to discriminate sounds on the basis of harmonic structure may be common to all vertebrates. Our results with tamarins provide additional support for this conclusion.

One alternative interpretation of our results is that tamarins are only sensitive to extremely drastic acoustic manipulations. This is evidenced by the fact that the tamarins responded more to calls that contained energy restricted to the fundamental or when the second harmonic was mistuned. Both of these conditions represent test calls with cross-correlation values lower than those of the missing second harmonic and missing fundamental conditions. However, this interpretation must be treated cautiously given the results from previous experiments (Weiss et al. 2001) in which the natural CLCs of some individuals in our colony could not be discriminated, at least given the design of our playback technique. The cross-correlation values obtained for those individuals whose calls could not be discriminated were lower than those reported for the fundamental alone and mistuned synthetics.

One limitation of our study is that it is difficult to interpret conditions in which the proportion of subjects responding to the test is intermediate between baseline and change of identity conditions reported previously (Weiss et al. 2001). This problem might be alleviated by increasing the sample size or quantifying a different response assay (e.g. duration and reaction time of response), which would facilitate other statistical techniques to assess the significance of the responses. Regardless, it is clear that most subjects do not attend to the fundamental frequency or second harmonic in determining individual identity. The majority of subjects treated calls without these components as if they were perceptually similar to the calls played back in the habituation series. It is possible though, that these elements are useful for other types of recognition, since our experimental context tested only individual recognition.

Research conducted with songbirds suggests that when an individual's own song is played back, the response elicited is intermediate between the types of responses given to familiar neighbours and those given to strangers (e.g. Brooks & Falls 1975; Searcy et al. 1981; McArthur 1986). Our study provided an additional opportunity to test this phenomenon with tamarins. Analyses of trials to habituation revealed no change in the number of trials needed for habituation in conditions in which subjects

were played their own calls as opposed to another individual's calls. Similarly, there was no significant difference in response to subject's own synthetic calls versus the synthetic calls of other individuals. These results are consistent with the findings from previous playback experiments using this colony of tamarins (Weiss et al. 2001). It is also consistent with results reported by Snowdon and colleagues (1983) in which cottontop tamarins respond to playbacks of their own quiet and normal long calls in the same way that they respond to a cagemate's calls. Furthermore, results from experiments designed to elicit antiphonal calling found no difference in the rate of antiphonal calling when subjects heard their own versus other's calls (Ghazanfar et al. 2002).

An additional finding from these experiments is that tamarin calls can effectively be synthesized and used to explore the mechanisms underlying acoustic perception. The synthetic exemplars were perceived as equivalent to natural exemplars in cases where subjects were habituated to calls from one individual and then played back a synthetic test call whose acoustics were derived from that same individual. In addition, in change of identity conditions, subjects responded to the synthetic in a manner that was similar to their response to natural exemplars. Finally, the high cross-correlation values for our synthetic exemplars indicate that this call type is highly amenable to synthesis.

The significance of our experiments is that they represent the beginning of a research programme whose aim is to identify the meaningful units of perception for cottontop tamarin vocalizations specifically, and nonhuman primate vocal repertoires more generally. Future experiments will continue to explore the effects of manipulating spectral and temporal parameters, using the detailed acoustic analyses of the tamarin CLC (Weiss et al. 2001), as well as a suite of different perceptual testing paradigms (Ghazanfar et al. 2001; Miller et al. 2001a, b).

### Acknowledgments

The authors would like to thank Kerry Jordan for her help in running the experiments and scoring trials. We would also like to thank Kim Beeman for his help in call synthesis. Thanks also to Asif Ghazanfar for comments on an earlier version of the manuscript, as well as to two anonymous referees of *Animal Behaviour*. This work was supported by Harvard University's Mind Brain and Behavior grant (368 422). The research presented here was approved by the Institutional Animal Care and Use Committee of Harvard University (Protocol No. 92–16).

### References

- Beeman, K. 1996. *Signal/RTS User's Guide*. Belmont, Massachusetts: Engineering Design.
- Boughman, J. W. 1997. Greater spear-nosed bats give group-distinctive calls. *Behavioral Ecology and Sociobiology*, **49**, 61–70.
- Bradbury, J. W. & Vehrencamp, S. L. 1998. *Principles of Animal Communication*. Sunderland, Massachusetts: Sinauer.
- Brooks, R. J. & Falls, J. B. 1975. Individual recognition in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. *Canadian Journal of Zoology*, **53**, 879–888.
- Burmeister, S., Wilczynski, W. & Ryan, M. J. 1999. Temporal call changes and prior experience affect graded signalling in the cricket frog. *Animal Behaviour*, **57**, 611–618.
- Cheney, D. L. & Seyfarth, R. M. 1980. Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, **28**, 362–367.
- Cheney, D. L. & Seyfarth, R. M. 1982. How vervet monkeys perceive their grunts: field playback experiments. *Animal Behaviour*, **30**, 739–751.
- Cheney, D. L. & Seyfarth, R. M. 1984. The acoustic features of vervet monkey grunts. *Journal of the Acoustical Society of America*, **75**, 1623–1628.
- 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.
- von Dommelen, W. 1990. Acoustic parameters in human speaker recognition. *Language and Speech*, **33**, 259–272.
- Emlen, S. T. 1972. An experimental analysis of the parameters of bird song eliciting species recognition. *Behaviour*, **41**, 130–171.
- Farabaugh, S. M., Brown, E. D. & Dooling, R. J. 1992. Analysis of warble song of the budgerigar *Melopsittacus undulatus*. *Bioacoustics*, **4**, 111–130.
- Farris, H. E., Forrest, T. G. & Hoy, R. R. 1997. The effects of calling song spacing and intensity on the attraction of flying crickets (*Orthoptera: Goryllidae: Nemobiinae*). *Journal of Insect Behavior*, **10**, 639–653.
- Fitch, W. T. & Kelley, J. P. 2000. Perception of vocal tract resonances by whooping cranes *Gras americana*. *Ethology*, **106**, 559–574.
- Garibaldi, B. T. 1999. Long calling in cotton-top tamarins. Senior thesis: Harvard University, Cambridge, Massachusetts.
- 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. & Watson, G. F. 1995. Within-male variability in call properties and female preference in the grey treefrog. *Animal Behaviour*, **50**, 1187–1191.
- Ghazanfar, A. A., Flombaum, J. I., Miller, C. T. & Hauser, M. D. 2001. The units of perception in cotton-top tamarin (*Saguinus oedipus*) long calls. *Journal of Comparative Physiology A*, **187**, 27–35.
- Ghazanfar, A. A., Smith-Rohrberg, D., Pollen, A. & Hauser, M. D. In press. Temporal cues in the antiphonal long calling behaviour of cottontop tamarins. *Animal Behaviour* **64** (200) UIS.
- Hammerschmidt, K. & Fischer, J. 1998. Maternal discrimination of offspring vocalizations in barbary macaques (*Macaca sylvanus*). *Primates*, **39**, 231–236.
- Handel, S. 1995. Timbre perception and auditory object identification. In: *Hearing*. 2nd edn. (Ed. by B. C. J. Moore), pp. 425–463. San Diego: Academic Press.
- Hartmann, W. M., McAdams, S. & Smith, B. K. 1990. Hearing a mistuned harmonic in an otherwise eriodic complex tone. *Journal of the Acoustical Society of America*, **88**, 1712–1724.
- Hauser, M. D. 1991. Sources of acoustic variation in rhesus macaque (*Macaca mulatta*) vocalizations. *Ethology*, **89**, 29–46.
- Hauser, M. D. 1996. *The Evolution of Communication*. Cambridge, Massachusetts: MIT Press.
- Jürgens, U. 1982. A neuroethological approach to the classification of vocalizations in the squirrel monkey. In: *Primate Communication* (Ed. by C. T. Snowdon, C. H. Brown & M. R. Petersen), pp. 50–62. Cambridge: Cambridge University Press.
- Kajikawa, S. & Hasegawa, T. 1996. How chimpanzees exchange information by pant-hoots: a playback experiment. *The Emergence of Human Cognition and Language*, **3**, 123–128.
- 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 Behavior Processes*, **26**, 261–273.
- 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.
- Leonard, M. L., Horn, A. G., Brown, C. R. & Fernandez, N. J.** 1997. Parent–offspring recognition in tree swallows, *Tachycineta bicolor*. *Animal Behaviour*, **54**, 1107–1116.
- Lohr, B. & Dooling, R. J.** 1998. Detection of changes in timbre and harmonicity in complex sounds by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, **112**, 36–47.
- McArthur, P.** 1986. Similarity of playback songs to self song as a determinant of response strength in song sparrows (*Melospiza melodia*). *Animal Behaviour*, **34**, 199–207.
- May, B., Moody, D. B. & Stebbins, W. C.** 1989. Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *Journal of the Acoustical Society of America*, **85**, 837–847.
- May, B. J., Moody, D. B. & Stebbins, W. C.** 1988. The significant features of Japanese monkey coo sounds: a psychophysical study. *Animal Behaviour*, **36**, 1432–1444.
- Miller, C. T., Dibble, E. & Hauser, M. D.** 2001a. Perception of an auditory illusion in a nonhuman primate. *Nature Neuroscience*, **4**, 783–784.
- Miller, C. T., Miller, J., Gil-da-Costa, R. & Hauser, M. D.** 2001b. Selective phonotaxis by cotton-top tamarins (*Saguinus oedipus*). *Behaviour*, **138**, 811–826.
- Moore, B. C. J.** 1989. *An Introduction to the Psychology of Hearing*. 3rd edn. San Diego: Academic Press.
- Moore, B. C. J., Glasberg, B. R. & Peters, R. W.** 1986. Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *Journal of the Acoustical Society of America*, **80**, 479–483.
- Owren, M. J.** 1990. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: II. Synthetic calls. *Journal of Comparative Psychology*, **104**, 29–40.
- Owren, M. J., Seyfarth, R. M. & Hopp, S. L.** 1997. The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*). *Journal of the Acoustical Society of America*, **101**, 2951–2963.
- Park, T. J. & Dooling, R. J.** 1986. Perception of degraded vocalizations by budgerigars (*Melopsittacus undulatus*). *Animal Learning & Behavior*, **14**, 359–364.
- Rendall, D., Rodman, P. S. & Edmond, R. E.** 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, **51**, 1007–1015.
- Sayigh, L. S., Tyack, P. L., Wells, R. S., Solow, A. R., Scott, M. D. & Irvine, A. B.** 1999. Individual recognition in wild bottlenose dolphins: a field test using playback experiments. *Animal Behaviour*, **57**, 41–50.
- Searcy, W. A., McArthur, P., Peters, S. S. & Marler, P.** 1981. Response of male song and swamp sparrows to neighbour, stranger and self songs. *Behaviour*, **77**, 152–163.
- Searcy, W. A., Nowicki, S. & Peters, S.** 1999. Song types as fundamental units in vocal repertoires. *Animal Behaviour*, **58**, 37–44.
- Simmons, A. M. & Bean, M. E.** 2000. Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana caatesbeiana*). *Journal of Comparative Psychology*, **114**, 167–173.
- 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**, 91–101.
- Weiss, D. J., Garibaldi, B. T. & Hauser, M. D.** 2001. The production and perception of long calls by cotton-top tamarins (*Saguinus oedipus*): acoustic analyses and playback experiments. *Journal of Comparative Psychology*, **15**, 258–271.