

Interruptibility of long call production in tamarins: implications for vocal control

Cory T. Miller^{1,*}, Stephen Flusberg² and Marc D. Hauser¹

¹Primate Cognitive Neuroscience Laboratory, Department of Psychology, Harvard University, Cambridge, MA 02138, USA and ²Department of Psychology, Northwestern University, Evanston, IL 60208, USA

*Author for correspondence (e-mail: ctmiller@wjh.harvard.edu)

Accepted 16 April 2003

Summary

Vocal production can be highly deterministic, such that once the central nervous system generates a signal to call, the vocalization is emitted immune to external events. Conversely, vocal production can be modulated by auditory feedback such that interference or disruption can cause an individual to stop calling or, if it continues to call, for the acoustic morphology of the signal to change. To explore which of these models best accounts for the control of vocal production in non-human primates, we adapted an interruption technique originally developed for songbirds for use with a New World monkey species, the cotton-top tamarin (*Saguinus oedipus*). Results from a pilot experiment indicated that an auditory stimulus (white noise) was more effective than a visual stimulus (strobe light) at interrupting the tamarin's species-typical 'combination long call (CLC)'. Data from a second

experiment showed that although the duration of the auditory stimulus did not affect the proportion of interruptions that occurred, a 1000 ms white noise stimulus perturbed the temporal structure of the CLC to a greater extent than did a 250 ms white noise stimulus. Furthermore, when call production was interrupted, tamarins stopped vocalizing after the completion of a syllable, suggesting that the syllable represents a unit of organization within the call. Overall, these results provide evidence that tamarins can modify their vocal output based on external events, but the degree of vocal control is significantly less than in oscine songbirds.

Key words: vocal production, vocal control, tamarin, *Saguinus oedipus*, combination long call, auditory stimulus, interruption.

Introduction

The mechanisms underlying vocal production have been explored in several species of non-human animals (see recent reviews in Hauser and Konishi, 1999). As with many aspects of non-human animal vocal behavior, there is extensive variation between species in the extent to which individuals can control vocal output (Nowicki, 1987; Nowicki and Capranica, 1986; Suthers and Fattu, 1973; Suthers et al., 1994). This behavioral variation can arise from a number of different factors, such as species-specific physiological constraints on the vocal production apparatus and/or limitations on auditory feedback. To achieve a deep understanding of the ultimate consequences and proximate causes of vocal control, a rich comparative data set is required.

An individual's control over vocal output is not simply an issue of presence or absence but rather a question of degree. At one end of the range of vocal control are species in which vocal production is entirely deterministic, with individuals showing no acoustic modification of vocal signals in response to external events. Species at the opposite end of the range will exhibit a capacity to modify individual elements within a vocalization in response to external events that occur during vocal production. Given the diverse acoustic structure of animal vocalizations,

however, it is often difficult to make meaningful cross-species comparisons of vocal control. One way to approach this issue, however, is to determine the extent to which different species can exert control over different components or features of a call. For species at the flexible end of the vocal control spectrum, individuals may have the capacity to control coarse-grained properties of the signal, such as overall duration and average pitch, as well as more fine-grained and individual components of the call, such as the number, arrangement and structure of syllables. By contrast, species at the more deterministic end of the range may show no capacity to exert control over components (e.g. syllables) of the signal. It should be noted that it is unlikely that any species will be either completely unlimited with regard to vocal plasticity or entirely immune to interference from external events. Experiments initially designed by Cynx (1990) for zebra finch (*Taeniopygia guttata*) and carried out by others working on different avian species provide a simple first step into approaching questions about the control of vocal signals. These experiments entail the playback of auditory or visual stimuli while an animal is vocalizing and asking whether sensitivity to this external event influences vocal output, perhaps by interrupting call production and/or

modifying the structure of the signal. In principle, this technique should be applicable to any vocalizing species, providing comparative data on vocal control.

Songbirds are known to possess dynamic abilities for vocal control, as evidenced by their capacity to engage in song matching with neighbors (Beecher et al., 1996, 2000), to show long-term changes in song structure in response to auditory perturbations (Leonardo and Konishi, 1999) and to simultaneously produce two independent acoustic signals (Greenewalt, 1968; Suthers, 1999). Given the multi-pulsed structure of birdsong, it is also important to assess the degree to which individuals are able to exert control over individual components within the song. For example, in Cynx's (1990) study of zebra finches, he flashed a strobe light while a bird emitted its species-specific song. Results showed that the strobe light caused zebra finches to interrupt singing 71% of the time. Typically, however, subjects completed the syllable already in production at the time of interruption before coming to a complete stop. Based on this finding, Cynx concluded that syllables must represent the smallest acoustic unit in the song at which vocal control can be exerted; no change in song structure could be induced on notes. Using comparable interruption techniques with other songbird (Reibel and Todt, 1997) and non-songbird (ten Cate and Ballintijn, 1996) species, results suggest that, at least for the avian vocal production system, individuals can exert control over individual syllables in a vocalization.

Traditionally, vocal production in non-human primates was thought to be highly deterministic, placing these species at the opposite end of the vocal control spectrum from songbirds and, more importantly, human primates (Lieberman, 1984). Support for this position is largely based on the paucity of evidence for ontogenetic change in call morphology (Hauser, 1989; Janik and Slater, 1997; Seyfarth and Cheney, 1986; Winter et al., 1973). Research indicated that the acoustic structure of infant and adult non-human primates was virtually identical, even when individuals were raised in acoustic and social isolation (Winter et al., 1973). However, a species' capacity for vocal learning may not be related to its ability to affect changes in a vocalization's structure in response to external events. In fact, despite limited evidence for vocal learning in non-human primates, several recent studies suggest that adults can modify the structure of their calls as a function of changes in the group's social dynamics. For example, studies of pygmy marmosets (*Cebuella pygmaea*; Elowson and Snowdon, 1994; Snowdon and Elowson, 1999), cotton-top tamarins (*Saguinus oedipus*; Weiss et al., 2001), Japanese macaques (*Macaca fuscata*; Sugiura, 1998) and chimpanzees (*Pan troglodytes*; Marshall et al., 1999; Mitani and Gros-Louis, 1998; Mitani and Brandt, 1994) demonstrate that individuals can produce vocalizations that converge on the acoustic structure of others in their group. Although it is not yet clear how the developmental and adult data are to be reconciled, the vocal convergence data at least suggest that some non-human primate species are equipped with the capacity to modify their own vocal output in response to a conspecific's vocalization. At present, the degree of vocal

control of non-human primates has yet to be fully determined. In parallel with studies on birds, an important first step in understanding non-human primate's capacity for vocal control is to examine whether vocal control can be exerted over individual components of the call.

We tested cotton-top tamarins in a paradigm similar to the one used by Cynx (1990) with zebra finches. Cotton-top tamarins produce a relatively large vocal repertoire consisting of discrete vocalizations containing either a single pulse or multi-syllabic units (Cleveland and Snowdon, 1981). In the following study, we focus on the 'combination long call' (CLC). The CLC is a highly stereotyped, multi-syllabic vocalization consisting of a concatenation of two acoustically distinct, and temporally discrete, syllable types: 1–2 chirps followed by 2–4 whistles (see Fig. 1; Cleveland and Snowdon, 1981; Miller et al., 2002). This long-distance vocalization is typically emitted when individuals are isolated from group members and elicits antiphonal calls (Ghazanfar et al., 2001; Miller et al., 2001a) as well as orienting and approach behavior (Miller et al., 2001b; C. T. Miller, J. S. Scarl and M. D. Hauser, submitted) from conspecifics. As such, it appears to function as a contact call (Miller and Ghazanfar, 2002) and may also be used as a signal for mate assessment (Miller et al., 2001b; C. T. Miller, J. S. Scarl and M. D. Hauser, submitted).

The goal of this study was to test between two general models of vocal control in a non-human primate, focusing specifically on the tamarins' CLCs. In the first model, when animals produce a vocal signal, commands from the central nervous system to the periphery function like a tape player without a stop button. Once the command is initiated, the signal plays through to completion, independent of any auditory feedback (internal or external). In the second model, the caller is equipped with a mechanism for vocal control, ostensibly a stop button on the tape player, one that enables it to arrest the signal at different points during production; within this second model, one can further explore whether the caller can stop at any point in the signal or only at specific points, with the latter providing useful information about the acoustic and motor organization of the signal. Here, we provide an initial test of these alternatives by exploring whether tamarin CLC production is interrupted by an external stimulus. If tamarin production is arrested, we can then further explore where signal production stops within the sequence of syllables and the extent to which call morphology is perturbed.

As no prior research of this kind had been conducted on non-human primates, we had no *a priori* reason to believe that either a visual or auditory stimulus would be more likely to interrupt tamarins. The first experiment therefore represents a pilot study in which we explore this issue by comparing whether a strobe light or a burst of white noise induces higher proportions of interruptions during call production. Based on the results from this experiment, indicating a higher level of interruption with white noise, we conducted a second study specifically testing how a burst of white noise influences the global syllable structure and temporal patterning of CLCs during interruption.

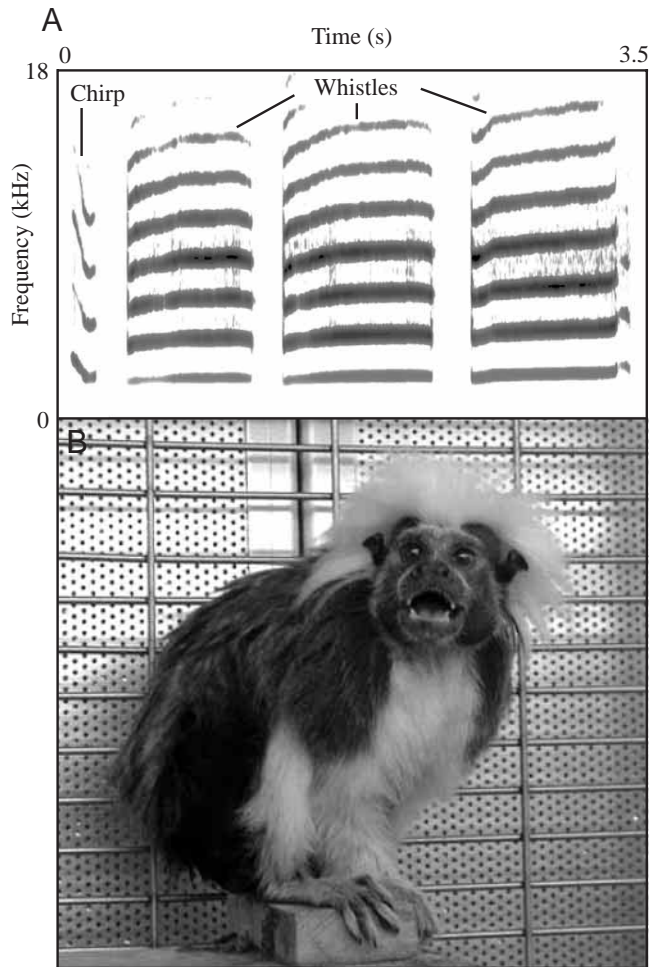


Fig. 1. (A) Spectrogram of a 'combination long call' with syllable types noted. (B) Cotton-top tamarin during vocal production.

EXPERIMENT 1: PILOT STUDY

Materials and methods

Subjects

Five adult cotton-top tamarins (*Saguinus oedipus* L.), two females and three males, participated as subjects in this experiment; we tested three males (AC, DD and RW) and two females (JG and RB) in the white noise condition, and one male (RW) and two females (JG and RB) in the strobe light condition. All subjects were born at the New England Regional Primate Research Center, Southborough, MA, USA and were housed at the Primate Cognitive Neuroscience Laboratory, Harvard University for the duration of the experiment. Their daily diet consisted of marmoset/tamarin chow (Purina, St Louis, MO, USA), crickets, fruit, peanuts and yogurt. They also had *ad libitum* access to water.

Apparatus

The apparatus used in this experiment was 20 cm×45 cm×60 cm. The sides and ceiling were constructed of opaque Plexiglas, while the front was wire mesh. The rear of the

apparatus was covered in white cloth and the floor was made of mesh wiring. A strobe light (xenon strobe; Realistic®, Fort Worth, TX, USA), used as a visual interruption stimulus, was placed 15 cm in front of the apparatus. White noise, used as the auditory interruption stimulus, was broadcast from an Alesis (Los Angeles, CA, USA) Monitor One speaker (Frequency Range 45–18000 Hz \pm 3 dB) positioned 1 m behind the apparatus. All trials were conducted in a sound-attenuated room.

General procedure

Subjects were removed from their home cage, carried to the experimental room inside a transport box and placed inside the testing apparatus. The experimenter then left the room and viewed the session from a separate room. Each time the subject initiated a long call the experimenter manually broadcast one of the two interruption stimuli. An attempt was made to initiate the stimulus as early as possible in the call. However, the brevity of the chirp made it difficult to broadcast the stimulus prior to the onset of whistle production. During trials involving visual interruption, the stimulus was presented while subjects were oriented towards the strobe light. For the entire duration of the experiment, subjects were in visual and auditory isolation from other tamarins and from the human experimenters. In cases when a subject remained in the test apparatus for over 2 min without calling, the experimenter broadcast the long call of another colony member in an attempt to elicit a long call from the test subject. This procedure was used because cotton-top tamarins typically produce CLCs in response to hearing a conspecific emit a CLC, a behavior known as antiphonal calling (Ghazanfar et al., 2001; Miller et al., 2001a). After each 10 min session, we returned the subject to its cage in the colony home room. Subjects participated in 2–4 sessions with the visual and/or auditory stimuli.

The duration of the strobe light was approximately 0.03–0.06 s. The white noise, generated in SoundEdit 16.2 (Macromedia, San Francisco, CA, USA), was 1000 ms in duration. We broadcast white noise at 68–70 dB sound pressure level (SPL) at 1 m from the speaker. Tamarins typically emit calls at approximately 60–65 dB SPL at the same distance. Thus, the interruption stimulus was louder than a typical CLC.

Data analysis

We recorded each session using a Video Labs camera and a Sennheiser microphone (Model ME-80). An 'interruption trial' was defined as all occasions when a subject emitted a long call and an interruption stimulus was broadcast during some portion of the call. We acquired all interruption trials at 30 frames per second onto a Macintosh G3 computer with Adobe Premiere software at a sampling rate of 44.1 kHz. We acquired all trials onto the computer and the sound file exported for analysis using CANARY (version 1.2; Cornell University, Ithaca, NY, USA). For each trial, we recorded the total number of chirps and whistles produced during interruption trials. In addition, we noted whether the interruption stimulus was broadcast during the production of a

syllable or during an inter-pulse interval (IPI). Furthermore, we measured the latency from the onset of the call to the onset of the interruption stimulus. We operationally defined ‘interrupted calls’ as any long call that terminated after the production of two or fewer whistle syllables (see Fig. 2). Previous acoustic analyses (Weiss et al., 2001) showed that the mean and mode number of whistles produced by cotton-top tamarins is three. As such, we assumed that any CLC consisting of fewer than three whistles was likely to be interrupted. All interruption stimuli in the Pilot Experiment were broadcast during the whistle portion of the CLC. The primary reason for the delay in stimulus onset was the hardware used in the experiment; this issue was rectified in Experiment 2.

Results

We recorded a total of 132 trials in which an interruption stimulus was broadcast while subjects emitted CLCs. Overall, tamarins interrupted vocal production in only 23% of trials. The two different types of interruption stimuli, however, were not equally successful at arresting long call production. While the strobe light interrupted calls 7% of the time (four out of 54 instances), the white noise stimulus interrupted calls 25% of the time (19 out of 77 instances). To test whether the proportion of interrupted calls elicited by these two stimulus types was significantly different, we calculated the percentage of interruptions for each subject and compared the proportion of interruptions between the two stimulus types in a factorial analysis of variance (ANOVA). Results indicated that subjects were more likely to be interrupted by the white noise stimulus than the strobe light ($F_{1,7}=5.53$; $P=0.05$). As all of the same subjects were not used in both test conditions, it was not possible to make within-subjects comparisons across these conditions.

Discussion

In this first, pilot experiment, we asked whether tamarin CLC production is interruptible on a general level and, more specifically, if the modality (vision *vs* audition) of the interrupting stimulus differentially affects the probability of interruption. Results indicated that tamarin CLCs can be interrupted by an external stimulus and that the probability of interruption is higher for white noise than for the strobe light. Building on these results, we conducted a second set of experiments using white noise stimuli of different durations to examine in greater detail the effects of the interruption stimulus on call structure. As the visual stimulus used in the pilot experiment was significantly shorter in duration than the white noise burst, it is possible that the increased call interruptions observed for the white noise stimulus resulted from a difference in stimulus duration rather than modality. To test this possibility, in Experiment 2 we broadcast white noise stimuli of 1000 ms (Condition A) and 250 ms (Condition B) in duration during CLC production.

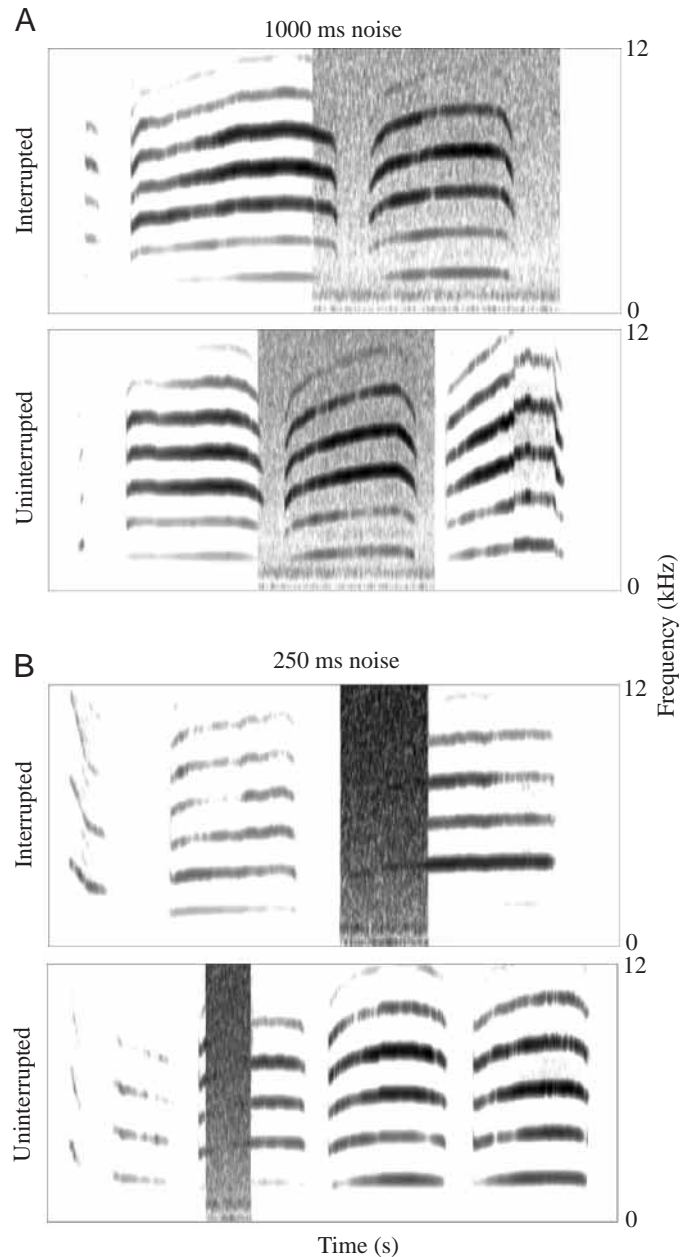


Fig. 2. Examples of interrupted and uninterrupted calls with (A) the 1000 ms white noise stimulus and (B) the 250 ms white noise stimulus. It should be noted that the *x*-axis for each of the spectrograms is on a different time scale because the duration of each call was different. A typical long call consisting of three whistles is approximately 3–3.5 s in duration. Only examples of calls in which the interruption stimulus was broadcast during the whistle portion of the call are shown. For some calls in Experiment 2, the noise stimulus was broadcast during the chirp portion of the vocalization.

The methodology employed in Experiment 2 differed from that used in the pilot experiment in three ways. First, rather than broadcast the stimulus during spontaneous and antiphonal call production, we focused on spontaneously produced CLCs. The reason for this change is because the mechanisms underlying spontaneous and antiphonal calling may differ and

affect flexibility in vocal control. Although comparing these two types of vocal production is an interesting question, we chose to focus on just one here. Second, during the first experiment, the stimulus was always broadcast during the whistle portion of the call. In Experiment 2, we broadcast the interruption stimulus during the chirp portion of the call when possible. As these two syllable types occur in stereotyped positions within the CLC, broadcasting the stimulus earlier in the call enabled us to assess whether the chirp and whistle portions of the call are represented as separate production units by tamarins and whether potentially interfering external events are more or less likely to cause tamarins to stop calling when they occur early or late in the signal. Third, rather than rely on population measures for the typical number of whistle units within the call to determine whether a call produced was interrupted, we ran a baseline session with each subject prior to testing in order to determine each subject's individual-specific call structure; more specifically, we determined the number of chirps and whistles for each subject during baseline sessions for comparison with the experimental conditions.

EXPERIMENT 2

Materials and methods

Subjects

Four adult cotton-top tamarins (two male: DD and RW; two female: RB and JG) participated as subjects in both conditions of this experiment. All subjects had participated in Experiment 1 and were the most consistent spontaneous callers in our colony.

Apparatus

We used the same apparatus as in Experiment 1, except that the speaker was positioned 1 m above the apparatus. A Sennheiser directional microphone (Model ME-80) was positioned 1 m in front of the speaker and aimed directly at the apparatus and all sessions recorded directly to a Tascam digital audiotape (DAT) recorder. Thus, the directional microphone aimed at subjects, rather than the speaker, enabled us to obtain a higher signal-to-noise ratio with respect to the vocalization and white noise. The apparatus was situated in a sound-attenuated booth (Model 400-A; Industrial Acoustics, Bronx, NY, USA) and all subject's behavior observed from a monitor outside the booth.

General procedure

Baseline session

We transported subjects from their home cage to the testing room in a transport box. Subjects were placed inside the test apparatus. The door to the sound-attenuated booth was closed and the session initiated. We recorded the first 10 spontaneously produced CLCs. All calls were recorded directly to DAT. The session lasted approximately 10 min. A baseline session was conducted prior to beginning the first test session of each condition.

Test session

The procedure used here was similar to that used in Experiment 1. Briefly, subjects were transported from the home cage to the testing room and placed in the test apparatus. Once subjects were in the apparatus, we closed the door of the sound-attenuated booth and began the test trial. The experimenter broadcast a burst of white noise each time a long call was spontaneously produced by the tamarin. In Condition A, we broadcast a 1000 ms white noise burst, while in Condition B we presented a 250 ms white noise burst. All white noise was generated using SoundEdit 16.2. We ran each subject on four test sessions for both Conditions A and B. However, during Condition B, three subjects (RB, DD and RW) were spontaneously producing calls at a low rate. As a result, each of these subjects was run on an additional test session to increase our sample of calls for the final analysis. We ran all subjects on Condition A before Condition B. For three of the subjects, the conditions were separated by 8 months, while for one subject (JG) only 3 weeks separated the different conditions.

As the stimuli differed in stimulus duration, it was necessary to control for the overall power of the stimulus. We broadcast the 1000 ms white noise stimulus at an intensity of 68–70 dB SPL and the 250 ms white noise stimulus at 88–90 dB SPL, measured 1 m from the speaker location. To ensure that the overall power was comparable between the two stimulus types, despite the difference in peak intensity, we recorded each of the stimuli to DAT, acquired the trial and measured the root mean square (RMS) amplitude over 1000 ms. The analysis of the 1000 ms stimulus involved only the duration of the stimulus, while the analysis of the 250 ms stimulus included both the stimulus as well as 750 ms of silence following the offset of the noise. This analysis revealed that the RMS amplitude over 1000 ms was approximately 3.28 μ Pa for both stimuli. Tamarins typically emit CLCs at 60–65 dB SPL at the same distance. Therefore, the intensity of the white noise was louder than a typical CLC.

Analysis

We acquired all test trials onto a Gateway computer using RTSD (Engineering Design, Belmont, MA, USA) at a sampling rate of 48 kHz. Using a customized signal analysis macro written for SIGNAL 3.1 (Engineering Design), we recorded the number of chirps and whistles for each call as well as the duration of each syllable and inter-pulse interval (IPI) for each CLC produced during baseline and test sessions. Following this analysis, we used CANARY (version 1.2) to extract values for each of the following parameters: location of interruption stimulus within the CLC, latency from the onset of call production to the interruption stimulus, latency between the onset of a syllable to the onset of the interruption stimulus for all calls in which the stimulus occurred within a syllable and the latency from the offset of the stimulus to the offset of the syllable in production for all calls in which the stimulus offset occurred within a syllable.

Table 1. Summary of temporal structure of baseline combination long calls (CLCs)

	Experimental subject			
	DD	RW	JG	RB
Condition A				
No. of whistles (mean±S.E.M.)	3.1±0.1	3.3±0.15	2.9±0.1	3.4±0.17
Min-max	3-4	3-4	2-3	3-4
Mode	3	3	3	3
No. of two-whistle calls	0	0	1	0
No. of three-/four-whistle calls	10	10	9	10
Duration of whistle (mean±S.E.M.)				
Whistle 1	446.2±49.3	717.1±59.6	725.4±38.8	404.2±31.6
Whistle 2	577.2±37.1	756.2±9.36	961.3±28.8	471.9±36.7
Whistle 3	572.2±27.9	681.9±14.1	764.5±19.3	467.6±47.7
Whistle 4	609.5±0	683.7±60.6	NA	436.1±106.9
Condition B				
No. of whistles (mean±S.E.M.)	3.2±0.2	3.1±0.2	2.8±0.13	3.5±0.2
Min-max	2-4	2-4	2-3	3-4
Mode	3	3	3	3
No. of two-whistle calls	1	1	2	0
No. of three-/four-whistle calls	9	9	8	10
Duration of whistle (mean±S.E.M.)				
Whistle 1	484.4±53.7	406.6±69.8	754.3±38.2	374.7±29.6
Whistle 2	667.8±44.1	653.1±61.1	972.1±46.9	527.4±34.6
Whistle 3	575.5±43.4	645.1±14.1	802.1±26.5	580.5±26.6
Whistle 4	523.5±65.6	577.8±11.2	NA	580.7±25.4

NA, not applicable.

Results

Baseline sessions

Data from the baseline trials conducted before both Condition A and Condition B revealed that the mean and mode number of whistle syllables for each individual subject was three, although subjects did occasionally produce calls consisting of two or four whistles (Table 1). As such, we considered all CLCs produced during the test conditions to be interrupted if they contained less than three whistle syllables. However, as subjects did on occasion produce CLCs consisting of two whistles (Table 1), it is possible that subjects would produce the same ratio of CLCs with less than three whistles in test conditions as in baselines. To test this, we conducted a repeated-measures ANOVA comparing the number of CLCs with less than three whistles and more than three whistles between baseline and test conditions. This analysis revealed that subjects were significantly more likely to produce CLCs with three or fewer whistles in test conditions than in baseline trials ($F_{1,3}=34.93$, $P=0.01$). Furthermore, there was no difference in this pattern between Conditions A and B ($F_{1,3}=0.34$, $P=0.6$). Overall, this suggests that although subjects do occasionally produce CLCs with two whistles, they are more likely to produce these calls following the broadcast of an interruption stimulus. It should be noted, however, that some proportion of the two-whistle CLCs produced during interruption trials may not have been interrupted, but rather were spontaneously produced two-whistle CLCs. The exact

proportion of interrupted calls cannot be accurately determined but, given the result here, that error is below the number of two-whistle CLCs produced during test trials. However, there is certainly some measurement error in the interruption percentages reported below.

Condition A

Patterns of CLC interruption

We broadcast a 1000 ms burst of white noise during 237 trials. Of these trials, 28% ($N=67$) were interrupted (see Table 2). Although the number of interruption trials varied between individuals due to variation in the propensity to spontaneously call, the percentage of interrupted calls was similar across subjects (Table 2). This proportion of interruption is comparable to the 25% of interruptions that were caused by the same stimulus in Experiment 1. Repeated-measures ANOVAs revealed that there were significantly more uninterrupted than interrupted calls ($F_{1,3}=44.51$, $P=0.007$) and no interaction between the pattern of interruptions and subjects ($F_{3,12}=0.32$, $P=0.81$), suggesting that the likelihood of interrupting call production was the same for all subjects. Similarly, there was no interaction between subject and session number ($F_{3,9}=0.27$, $P=0.85$). To test whether subjects showed any changes in the propensity to interrupt calling, we conducted a series of regressions that plotted proportion of interruptions by test session for each subject. Results indicated that two subjects showed no change in call interruptibility (DD:

Table 2. *Interruptibility of long calls*

Subject	No. of interrupted calls		Total	% Interrupted
	Yes	No		
Condition A				
JG	16	37	53	30
RB	24	53	77	31
RW	6	24	30	25
DD	21	56	77	38
Total	67	170	237	28
Condition B				
JG	8	60	68	11
RB	18	40	58	31
RW	8	47	55	15
DD	17	30	47	36
Total	51	177	228	22

$r^2=0.45$, $P=0.33$; JG: $r^2=0.11$, $P=0.66$), while two subjects showed a significant decrease in the proportion of calls interrupted over successive test sessions (RB: $r^2=0.97$, $P=0.005$; RW: $r^2=0.98$, $P=0.009$). This suggests that at least some tamarins may be able to adjust vocal production in response to consistent perturbations in the environment.

Given the stereotyped syllable order of CLCs, it is possible that the chirp and whistle portions of the call represent distinct production units in the CLC. To determine whether the interrupting stimulus was more likely to elicit an interruption depending on whether it was broadcast during the chirp or whistle portion of the call, we analyzed whether the trials in which the stimulus occurred during the chirp portion of the call differed from the overall pattern. A total of 29 interruption stimuli were broadcast during the chirp portion of the CLC. Of these calls, 10 were interrupted (34%). The caller stopped emitting the call before producing any whistle syllables in three cases. In the remaining seven trials, subjects produced one or two whistles before interrupting call production. This represents a comparable proportion of interruptions to the overall pattern for this condition (28%).

When a call is experimentally interrupted, the amount of signal that follows the interruption stimulus may provide one measure of the organization of acoustic units and the degree of vocal control that can be exerted over these units. To examine this, we analyzed the relationship between when the interruption stimulus was broadcast and when the call was interrupted. Results indicated that during interrupted calls, subjects never immediately stopped calling but rather continued call production until at least the end of the syllable. In fact, most callers produced an additional syllable before call interruption. The typical pattern was for the interruption stimulus to be broadcast during either the chirp portion or first whistle of the CLC. Rather than immediately interrupt call production, subjects continued calling through two whistle syllables. In other words, the general result was for callers to produce CLCs consisting of two whistles independent of when

the white noise was broadcast. Of all interrupted calls, three contained zero whistles, nine contained one whistle and 55 contained two whistles. The propensity to produce CLCs consisting of two whistles is evident when we compare the location of the stimulus onset and the timing of call interruption. In 40 of the 67 interrupted calls, the interruption stimulus was broadcast prior to the production of the second whistle, but callers continued call production until the completion of the second whistle (see Fig. 2).

Next, we compared the proportion of interrupted calls for trials in which the stimulus was broadcast during a syllable or an IPI. Overall, we broadcast white noise during the middle of a syllable (chirp or whistle) on 184 trials and, of these, 49 (38%) were interrupted. By contrast, we broadcast noise during an IPI on 53 trials and, of these, 18 were interrupted (34%). A factorial ANOVA revealed no difference in the likelihood of interrupting a caller whether the stimulus was broadcast either during syllable production or an IPI ($F_{1,12}=1.69$, $P=0.22$).

The following set of analyses tested whether the timing of the stimulus affected the likelihood of eliciting an interruption for the following three latency measures. First, we wanted to determine whether the latency from the onset of the CLC to the onset of the stimulus affected interruptibility of the call. Analyses revealed no statistically significant difference in this measure between interrupted (mean \pm S.E.M., 671.0 ± 39.2 ms) and uninterrupted calls (676.1 ± 23.7 ms; $t_{235}=0.1$, $P=0.92$). Second, we conducted a similar analysis as the first but focused on the timing of the stimulus within a syllable for both chirps and whistles. Here, we examined whether the latency from the onset of syllable production to the onset of the white noise affected call interruption for all trials in which the stimulus onset occurred within the syllable. The latency between interrupted (212.3 ± 27.9 ms; $N=42$) and uninterrupted (272.9 ± 19.8 ms; $N=113$) calls on this measure was not significantly different ($t_{153}=0.65$, $P=0.1$). Third, we analyzed whether the latency from the offset of the stimulus to the offset of the syllable in production influenced interruption of call production for calls in which the interruption stimulus ended within a syllable. This result essentially mirrors the previous analysis. Again, results indicated no difference in this latency measure for interrupted (329 ± 39.6 ms; $N=18$) and uninterrupted (305.5 ± 15.2 ms; $N=129$) calls ($t_{145}=0.55$, $P=0.58$).

Acoustic modification of CLCs

To determine whether the temporal structure of CLCs was affected by the interruption stimulus, we compared the acoustic structure of CLCs produced during test trials to calls from baseline trials. We focused our analyses on Whistle 1, 2 and the IPI between these whistles because the interruption stimulus overlapped with this portion of nearly all calls produced during interruption trials. All such analyses were conducted using unpaired *t*-tests (Table 3). Overall, calls produced during the baseline sessions had significantly longer ISIs than during test trials. It is possible that this difference resulted because in interrupted calls the measured ISI was the final one of the call, but for the baseline calls the ISI was

interrupted and uninterrupted calls. Analyses revealed a significant difference between the latency from the onset of the CLC to the onset of the stimulus for interrupted (845.5 ± 23.87 ms) and uninterrupted (941.1 ± 39.4 ms) calls ($t_{226}=1.94$, $P=0.05$). Second, we analyzed whether the latency from the onset of the syllable in production to the onset of the stimulus differed between interrupted and uninterrupted calls. There was no statistically significant difference between interrupted (358.8 ± 20.60 ms; $N=46$) and uninterrupted (333.5 ± 29.4 ms; $N=144$) calls ($t_{188}=0.63$, $P=0.53$). Third, the latency from the offset of the stimulus to the offset of the syllable in production during the stimulus offset was not significantly different for interrupted (352.8 ± 55.6 ms; $N=24$) as opposed to uninterrupted (386.9 ± 18.3 ms; $N=148$) calls ($t_{170}=0.68$, $P=0.50$).

Acoustic modification of CLCs

Our final analysis focused on how the interruption stimulus affected the acoustic structure of the CLC. As in the previous condition, this analysis was restricted to the temporal measure of the first two whistles and the ISI between these whistles. As shown in Table 3, no single feature differed significantly across all subjects as a result of the interruption stimulus. All but one subject did, however, show a significant change in at least one acoustic parameter.

Comparison of Conditions A and B

Overall, there appears to be little difference between Conditions A and B. In Condition A, all subjects were interrupted at roughly comparable proportions, while two subjects exhibited a noticeably lower proportion of interrupted calls in Condition B relative to Condition A. However, a repeated-measures ANOVA revealed that there was no significant difference in the overall proportion of interrupted calls between the two conditions ($F_{1,3}=0.24$, $P=0.66$). Similarly, when we included subject and condition as between-subject variables, the three-way interaction between subject, condition and interruption pattern was not significant ($F_{3,28}=1.71$, $P=0.19$).

As two subjects showed evidence of habituation in Condition A, we wanted to determine whether there was a difference in the patterns of interruption between the first sessions of both Conditions A and B. Results indicated no interaction between call interruptions and test session ($F_{1,3}=0.67$, $P=0.473$). Similarly, there was no difference in the proportions of interruptions between the final test session in Condition A and the first test session in Condition B ($F_{1,3}=2.49$, $P=0.21$). These analyses suggest that no habituation to the interruption stimulus occurred between test conditions.

Discussion

In the introduction, we described two general models of vocal control in animals. Model 1 presumes that when animals call, a signal from the central nervous system initiates a command to the motor system, which then initiates call

production; the signal plays out, independent of external stimuli. Model 2 presumes that callers have some control over production. When the central nervous system sends a command, the caller initiates the call but can stop or modify call structure depending upon the nature of external stimuli; the timing of call cessation may provide some information about both motor control and the organization of the call's acoustic morphology, specifically its production units. As it is unlikely that any species exists that represents the extremes of either of these two systems, it is necessary to test each species' vocal plasticity to determine where they reside on this control spectrum. To explore which of these models best characterizes the system of call production in cotton-top tamarins, we used the interruption paradigm originally designed and implemented with birds (Cynx, 1990). The logic of this technique is that if subjects stop calling before the normal end point in the call, then the point of cessation informs our understanding of acoustic organization, and provides evidence for production units below that of the whole call. If, however, subjects continue to call through the interruption stimulus, with no evidence of modifying call structure from baseline conditions, then it suggests that the caller has no control over individual components of a call. Although no previous studies of this kind have been attempted in non-human primates, the tamarin represents an ideal candidate for such an experiment due to prior knowledge of its vocal repertoire, together with the multipulsed structure of the CLC.

As the goal of this study was to determine how external events influence call production, it was important to ascertain what constitutes a normal CLC. Previous acoustic analyses showed that typical CLCs consist of a single chirp followed by three whistles (Weiss et al., 2001); however, some variability in call structure does exist. To address this issue, we conducted a series of baseline calling sessions and measured the number of chirps and whistles produced in each call. Based on these sessions, we asserted that any call consisting of fewer than three whistles could be considered interrupted. Statistical analyses showed that subjects were more likely to produce calls consisting of fewer than three whistles during test trials compared with baseline sessions, suggesting that this definition of interruption is valid. Rather than assert that our metric of interruption is definitive, we consider our notion of what constitutes an interrupted call to be a working definition. For the purpose of this study, our definition was sufficient, but it is possible that future studies will find that a refined definition is more appropriate.

The first experiment, a pilot study, tested whether an auditory stimulus (white noise) or a visual stimulus (light flash) would elicit higher proportions of call interruptions in tamarins. Results indicated that, unlike birds, the light flash had little effect on call production, causing interruption on only 7% of the trials, while the white noise stimulus interrupted subjects on 25% of trials. This suggested that an auditory stimulus would be more effective as an interruption stimulus than a visual stimulus. Building on these results, we conducted a second set of experiments to look more closely at how vocal

production and the acoustic structure of the CLC were affected by an interruption stimulus.

Experiment 2 consisted of two experimental conditions that differed only in the duration of the interruption stimulus. The white noise stimulus for Condition A was 1000 ms in duration, while a 250 ms noise stimulus was used in Condition B. We used a baseline calling condition to determine that our subjects typically produced three whistle syllables during the production of CLCs. Thus, we assumed that any call consisting of fewer syllables had been interrupted. Overall, we observed that both of these stimuli were equally likely to interrupt vocal production (A: 28%; B: 22%). In fact, there was no statistically reliable difference in the interruption patterns of subjects across these two experimental conditions, suggesting that although the proportion of interrupted calls is lower than in some studies with birds using visual stimuli (i.e. Cynx, 1990; Reibel and Todt, 1997; ten Cate and Ballintijn, 1996) the effect is stable and repeatable in tamarins.

In contrast to birds, tamarins did not interrupt vocal production immediately after completing the syllable already in production at the onset of the interruption stimulus. Instead, the typical pattern of interruption was for callers to produce two whistle syllables independent of when the interruption stimulus was broadcast. This pattern emerged despite the fact that for the majority of calls the onset of the stimulus occurred prior to the onset of the second whistle (i.e. during the chirp section or first whistle). At least three explanations exist for this effect. First, the introductory chirp(s) and first two whistles represent a functionally significant stable production unit. In a recent study of tamarin antiphonal calling behavior, Ghazanfar et al. (2002) observed that, following the playback of a CLC, subjects did not emit their vocal response until at least three syllables had played, typically one chirp and two whistles. It may be that two whistle syllables represent a threshold for the minimum amount of acoustic information necessary for call recognition. As a result, the vocal production system may have evolved to complement this threshold. Second, although vocal production was arrested before the expected completion of the call, the call was not strictly interrupted. Rather than cause a complete halt in vocal production, the stimulus induced the tamarin auditory feedback mechanism to attempt to compensate for the noise. The longer stimulus duration in Experiment 2, Condition A induced longer ISIs for all subjects, suggesting that modification of call structure can occur during vocal production. The consistent production of two whistles following stimulus presentation may represent evidence of vocal modification rather than call interruption. Third, it may be that the mechanism that signals the vocal production system to cease calling cannot initiate vocal arrest immediately but rather is delayed. Here, the latency in call interruption is entirely due to constraints imposed by the mechanisms underlying the auditory feedback system. At present, it is not possible to determine which of these possibilities causes tamarins to typically arrest calling after only two whistles have been produced. Future experiments, however, will address this issue.

The overall proportion of interrupted calls between Conditions A and B did not differ significantly. However, the duration of the interruption stimulus did seem to affect the temporal acoustic structure of CLCs (see Table 3). In Condition A, the 1000 ms stimulus elicited significant changes in the IPI between the first two whistles, contrasting CLCs produced during test and baseline trials across all subjects. Specifically, the duration of the IPI increased for all test trials compared with baselines. In contrast to the effects demonstrated in Condition A, the 250 ms stimulus used in Condition B only affected CLC structure on an individual level. Namely, most subjects produced calls that differed in their acoustic structure as a result of the interruption stimulus, but which aspects of the call changed varied for each subject. Although the only difference between the experimental conditions in Experiment 2 was the duration of the stimulus used, results indicated that the 1000 ms stimulus affected the acoustic structure of CLCs more consistently than did the 250 ms stimulus. This difference may be explained in terms of the auditory feedback necessary for the maintenance of the CLC acoustic structure.

Although most primate species are thought to only exhibit a limited degree of vocal control, humans are an exception. Much like songbirds, humans exhibit extensive control of vocal output and specialized mechanisms for vocal learning (Jusczyk, 1997; Locke, 1993; Pinker, 1994; Stevens, 1998; Titze, 1994). Although humans are part of the primate order, their capacity for vocal control represents an outlier; they are much more like songbirds and some cetaceans. As a result, it is interesting to ask whether the pattern of interruptibility in humans is more like non-human primates or songbirds. Data from an experiment conducted by Ladefoged et al. (1973) may provide some data that address this question. In these experiments, experimenters broadcast a tone while subjects produced a memorized sentence, and the latency to stop vocal production was measured. Results indicated that subjects typically ceased vocal production approximately 200–300 ms following the onset of the interruption stimulus. However, if experimenters broadcast the interruption stimulus immediately before subjects began a vocal utterance, the latency was significantly longer. Although the procedure employed in this experiment is not directly comparable with the present study or experiments on birds, these data suggest that, like songbirds, human vocal production can be quickly stopped in response to external events.

Auditory feedback is necessary for many species of songbirds to maintain the stereotyped structure of their species-typical song (Brainard and Doupe, 2000; Cynx and Von Rad, 2001; Leonardo and Konishi, 1999). However, there is a noticeable absence of data on this topic in non-human primates. The paucity of studies in this area is probably due to a general consensus that non-human primates lack the ability for vocal learning during both ontogeny and adulthood (Janik and Slater, 1997; Seyfarth and Cheney, 1997). The rationale here is that if non-human primates are incapable of modifying the acoustic structure of their vocalizations then studies of auditory feedback

are unnecessary. However, recent studies suggest that adult apes (Marshall et al., 1999; Mitani and Gros-Louis, 1998; Mitani and Brandt, 1994), Old World monkeys (Suguiwa, 1998) and New World monkeys (Elowson and Snowdon, 1994; Snowdon and Elowson, 1999; Weiss et al., 2001) are capable of modifying call structure, allowing for convergence within group members. The study reported here, combined with the aforementioned studies, suggests that the malleability of non-human primate vocal production may be greater than was previously thought. Rather than simply uttering innate, stereotyped vocalizations, non-human primates seem able to exert some vocal control over their vocalizations and affect changes in the call's structure during vocal production. The differences in the results obtained for vocal learning during ontogeny and vocal flexibility in adulthood may suggest that different mechanisms underlie these vocal behaviors. Future work will need to explore this issue in more detail.

We thank Roian Egnor, Asif Ghazanfar and two anonymous referees for helpful comments on this manuscript. This research was supported by an NIH predoctoral individual NRSA fellowship to C.T.M. (F31 MH63501) and grants from the NSF (LIS-411028-G; ROLE) and Harvard University to M.D.H. Some members of the tamarin colony were originally provided by New England Regional Primate Research Center, Southborough, MA. All research described here was approved by the Harvard University Animal Care and Use committee (A3598-01; 92-16).

References

Beecher, M. D., Campbell, S. E., Burt, J. M., Hill, C. E. and Nordby, J. C. (2000). Song type matching between neighbouring song sparrows. *Anim. Behav.* **59**, 21-27.

Beecher, M. D., Stoddard, P. K., Campbell, S. E. and Horning, C. L. (1996). Repertoire matching between neighboring song sparrows. *Anim. Behav.* **51**, 917-923.

Brainard, M. S. and Doupe, A. J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**, 762-766.

Cleveland, J. and Snowdon, C. T. (1981). The complex vocal repertoire of the adult cotton-top tamarin, *Saguinus oedipus*. *Zeitschrift für Tierpsychologie* **58**, 231-270.

Cynx, J. (1990). Experimental determination of a unit of song production in the zebra finch (*Taeniopygia guttata*). **104**, 3-10.

Cynx, J. and Von Rad, U. (2001). Immediate and transitory effects of delayed auditory feedback on birdsong production. *Anim. Behav.* **62**, 305-312.

Elowson, A. M. and Snowdon, C. T. (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Anim. Behav.* **47**, 1267-1277.

Ghazanfar, A. A., Flombaum, J. I., Miller, C. T. and Hauser, M. D. (2001). The units of perception in cotton-top tamarin (*Saguinus oedipus*) long calls. *J. Comp. Physiol. A* **187**, 27-35.

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

Greenewalt, C. H. (1968). *Bird Song: Acoustics and Physiology*. Washington, DC: Smithsonian Institution Press.

Hauser, M. D. (1989). Ontogenetic changes in the comprehension and production of vervet monkey (*Cercopithecus aethiops*) vocalizations. *J. Comp. Psychol.* **103**, 149-158.

Hauser, M. D. and Konishi, M. (1999). *The Design of Animal Communication*. Cambridge, MA: MIT Press.

Janik, V. M. and Slater, P. J. B. (1997). Vocal learning in mammals. *Adv. Study Behav.* **26**, 59-99.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge: MIT Press.

Ladefoged, P., Silverstein, R. and Papcun, G. (1973). Interruptibility of speech. *J. Acoust. Soc. Am.* **54**, 1105-1108.

Leonardo, A. and Konishi, M. (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* **399**, 466-470.

Lieberman, P. (1984). *The Biology and Evolution of Language*. Cambridge, MA: Harvard University Press.

Locke, J. (1993). *The Path to Spoken Language*. Cambridge, MA: Harvard University Press.

Marshall, A. J., Wrangham, R. W. and Clark, A. P. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Anim. Behav.* **58**, 825-830.

Miller, C. T., Dibble, E. and Hauser, M. D. (2001a). Amodal completion of acoustic signals by a nonhuman primate. *Nat. Neurosci.* **4**, 783-784.

Miller, C. T. and Ghazanfar, A. A. (2002). Meaningful acoustic units in nonhuman primate vocal behavior. In *The Cognitive Animal* (ed. M. Bekoff, C. Allen and G. Burghardt), pp. 265-273. Cambridge, MA: MIT Press.

Miller, C. T., Miller, J., Costa, R. G. D. and Hauser, M. D. (2001b). Selective phonotaxis by cotton-top tamarins (*Saguinus oedipus*). *Behaviour* **138**, 811-826.

Miller, C. T., Weiss, D. J. and Hauser, M. D. (2002). Mechanisms of acoustic perception in cotton-top tamarins. In *Primate Audition: Behavior and Neurobiology* (ed. A. A. Ghazanfar), pp. 43-60. New York: CRC Press.

Mitani, J. and Gros-Louis, J. (1998). Chorusing and convergence in chimpanzees: tests of three hypotheses. *Behaviour* **135**, 1041-1064.

Mitani, J. C. and Brandt, K. L. (1994). Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology* **96**, 233-252.

Nowicki, S. (1987). Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature* **325**, 53-55.

Nowicki, S. and Capranica, R. R. (1986). Bilateral syringeal coupling during phonation of a songbird. *J. Neurosci.* **6**, 3593-3610.

Pinker, S. (1994). *The Language Instinct*. New York: William Morrow and Co., Inc.

Reibel, K. and Todd, D. (1997). Light flash stimuli alters the nightingale's singing style: implications for song control mechanisms. *Behaviour* **134**, 789-808.

Seyfarth, R. M. and Cheney, D. L. (1986). Vocal development in vervet monkeys. *Anim. Behav.* **34**, 1640-1658.

Seyfarth, R. M. and Cheney, D. L. (1997). Some general features of vocal development in nonhuman primates. In *Social Influences on Vocal Development* (ed. C. T. Snowdon and M. Hausberger), pp. 249-273. Cambridge: Cambridge University Press.

Snowdon, C. T. and Elowson, A. M. (1999). Pygmy marmosets modify call structure when paired. *Ethology* **105**, 893-908.

Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Suguiwa, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Anim. Behav.* **55**, 673-687.

Suthers, R. A. (1999). The motor basis of vocal performance in songbirds. In *The Design of Animal Communication* (ed. M. D. Hauser and M. Konishi), pp. 651-665. Cambridge, MA: MIT Press.

Suthers, R. A. and Fattu, J. M. (1973). Mechanisms of sound production in echolocating bats. *Am. Zool.* **13**, 1215-1226.

Suthers, R. A., Goller, F. and Hartley, R. S. (1994). Motor dynamics of song production by mimic thrushes. *J. Neurobiol.* **25**, 917-936.

ten Cate, C. and Ballintijn, M. R. (1996). Dove coos and flashed lights: interruptibility of "song" in a nonsongbird. *J. Comp. Psychol.* **110**, 267-275.

Titze, I. R. (1994). *Principles of Voice Production*. Englewood Cliffs, NJ: Prentice Hall.

Weiss, D. J., Garibaldi, B. T. and Hauser, M. D. (2001). The production and perception of long calls by cotton-top tamarins (*Saguinus oedipus*): acoustic analyses and playback experiments. *J. Comp. Psychol.* **11**, 258-271.

Winter, P., Handley, P., Ploog, W. and Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour* **47**, 230-239.