



Discrete or graded variation within rhesus monkey screams? Psychophysical experiments on classification

C. G. LE PRELL*, M. D. HAUSER† & D. B. MOODY*

*Kresge Hearing Research Institute, Department of Otolaryngology, University of Michigan

†Department of Psychology and Program in Neurosciences, Harvard University

(Received 1 November 2000; initial acceptance 30 January 2001;
final acceptance 25 May 2001; MS. number: A8923)

Gouzoules et al. (1984, *Animal Behaviour*, 32, 182–193) presented evidence that semifree-ranging rhesus monkeys, *Macaca mulatta*, produce acoustically distinctive classes of scream vocalizations that carry different functional messages. To determine the perceptual validity of these vocal classes, we conducted psychophysical experiments on captive rhesus monkeys. We trained two monkeys to maintain contact with a metal response cylinder during presentation of nontarget stimuli, and to release the cylinder to report detection of target stimuli. For one subject, tonal screams served as nontarget stimuli and arched screams served as targets. These conditions were reversed for a second subject. Once natural exemplars were correctly discriminated, both subjects correctly generalized to synthetic targets. Variability in responses to nontarget stimuli, however, suggested that scream categories were not well defined following training. This result suggests that rhesus monkeys do not perceive categorical distinctions between arched and tonal screams, at least under the testing conditions implemented. Rather, our results provide evidence for a graded category. To explore which acoustic features are most important for classifying novel exemplars as tonal or arched screams, we ran several follow-up experiments with novel scream exemplars. Generalization trials suggested that variation in rate of frequency change, maximum frequency of the fundamental and harmonic structure may be important to the discrimination of screams.

© 2002 The Association for the Study of Animal Behaviour

Identifying acoustic features that define vocal signal classes has become increasingly important as evidence has accumulated that primates classify vocalizations based on their meaning (for review, see Hauser 1996). Investigations into meaningful signal classes can be based on correlations between vocalizations and socioecological context (e.g. Struhsaker 1967; Green 1975; Seyfarth et al. 1980a, b; Gouzoules et al. 1984; Hauser & Marler 1993), or demonstration that playback of different signals reliably elicits different responses (e.g. Seyfarth et al. 1980a, b; Gouzoules et al. 1984; Cheney et al. 1995; Zuberbühler et al. 1997; Fischer et al. 1998; Cheney & Seyfarth 1999; Rendall et al. 1999). Alternatively, a change in response due to a change in a repeating signal can be observed ('habituation/dishabituation', e.g. Cheney & Seyfarth 1988; Fischer 1998; Hauser 1998). In

laboratory-based experiments, animals can be isolated and vocal signal classification further examined (e.g. Petersen et al. 1978, 1984; Beecher et al. 1979; Brown et al. 1979; Zoloth et al. 1979; Owren, 1990a, b; Hopp et al. 1992). We applied laboratory-based methods to study the classification of scream calls by rhesus monkeys, *Macaca mulatta*. The vocalizations were drawn from 'arched' and 'tonal' scream classes. Field studies suggest these call types are acoustically and functionally distinct (Gouzoules et al. 1984), and more quantitative analyses indicate that comparable categories exist in other closely related macaque species (Gouzoules & Gouzoules 1989). Our experiments explore whether rhesus perceive arched and tonal screams as a discrete or graded class of vocalizations.

Observations of free-ranging rhesus monkeys on Cayo Santiago, Puerto Rico, suggest that rhesus produce scream vocalizations within the context of submission (Gouzoules et al. 1984; Hauser & Marler 1993; Rendall et al. 1998). Functionally, the acoustic morphology of screams covaries with opponent social ranking and severity of aggression directed at the caller. For example,

Correspondence: C. G. Le Prell, 4030 Kresge Hearing Research Institute, 1301 East Ann Street, University of Michigan Medical Campus, Ann Arbor, MI 48109-0506, U.S.A. (email: colleeng@umich.edu). M. D. Hauser is at the Department of Psychology and Program in Neurosciences, 33 Kirkland Street, Harvard University, Cambridge, MA 02138, U.S.A.

arched screams are primarily given by lower-ranking individuals in response to aggressive threats lacking physical contact. In contrast, tonal screams are primarily produced by subordinates in response to aggression by a relative, and sometimes in situations involving aggression from higher- or lower-ranking individuals in the absence of physical contact (see Gouzoules et al. 1984). Although the association of call classes with opponent attributes deviate from that predicted by chance alone, the calls do not sort perfectly by context. For example, noisy screams were produced more often than any other type of scream in virtually all of the aggressive/social contexts described by Gouzoules et al. (1984).

Because free-ranging female rhesus differentially aid other monkeys during agonistic encounters (e.g. Kaplan 1978), vocal signals could be used to elicit aid based on recognition of caller identity. Although adult female rhesus monkeys do not differentially respond to noisy screams produced by kin and nonkin (Rendall et al. 1998), Rendall et al. (1996) observed differential responsiveness to coos produced by kin and nonkin. Female rhesus also respond differentially to different types of screams produced by their own offspring (Gouzoules et al. 1984). Although response latencies in Gouzoules et al.'s study were greater than those reported by Rendall et al. (1996, 1998), female rhesus clearly respond differently to different classes of screams (Gouzoules et al. 1984). Thus, it seems likely that rhesus monkeys can use the acoustic morphology of screams to determine the type of opponent attacking a caller as well as the caller's identity. However, caller identity may be more difficult to extract from some screams than others.

Acoustically, arched screams are characterized by 'an initial sharply ascending frequency modulation forming one or two peaks, followed by a sharply descending frequency modulation', (Gouzoules et al. 1984). Tonal screams may contain similar beginning and ending components, but always contain 'a distinctive tonal mid-section', (Gouzoules et al. 1984). In addition to these classification criteria, tonal screams are significantly longer in duration, tend to be composed of lower-frequency components, and encompass a significantly narrower frequency bandwidth than do arched screams (Gouzoules et al. 1984). Although there is some overlap of acoustic characteristics across scream categories, Gouzoules et al. (1984) reported that approximately 90% of the calls they recorded could be classified into relatively discrete categories. Observations during collection of the present vocalization set revealed more gradation and heterogeneity of call bouts, however, suggesting a graded system. Indeed, other investigators have reported difficulty sorting these sounds into discrete categories suggesting instead that this class of vocalizations is heterogeneous (e.g. Kitko et al. 1999; M. J. Owren & M. Nederhouser, unpublished data). The following experiments were designed to determine whether rhesus monkeys form a discrete or graded representation of the acoustic variation in scream vocalizations.

GENERAL METHOD

Subjects

Two female rhesus macaques were obtained from the University of Florida (female m181) and the University of Southwest Louisiana (female m184). M184 was raised in a semifree-ranging environment, and was approximately 3.5 years old upon arrival at the University of Michigan. M181 was born at the University of Wisconsin (Wisconsin Regional Primate Center) and subsequently transferred to the University of Florida. While at the University of Florida, this animal was pair-housed and exercised in an outdoor enclosure. She was approximately 9 years old upon arrival at the University of Michigan. Neither subject had previously participated in a psychophysical investigation. Female subjects were selected because free-ranging rhesus females provide aid during agonistic situations more often and more effectively than males (Kaplan 1977, 1978).

Subjects were individually housed with free access to water, and maintained on slightly restricted diets. During operant sessions, subjects received 190 mg of whole-diet food pellets (Bio-Serv Precision Banana Pellets) for correct responses. After each session, daily food rations were completed with Purina High Protein Monkey Chow. Vitamins were provided daily and fresh fruit was provided several times each week. Foraging boards filled with hay, seeds and nuts were also provided at least once weekly (for reviews of environmental enrichment protocols, see Fajzi et al. 1989; Schapiro et al. 1991). Animal treatment met or exceeded all guidelines in the *Guide for the Care and Use of Laboratory Animals* (National Research Council 1996).

Training Procedures

We used positive reinforcement procedures to train subjects to maintain contact with a metal response cylinder during presentation of standard (i.e. nontarget) stimuli, and to release contact to report detection of target stimuli. Standard and target stimuli were screams drawn from arched and tonal scream categories. We completed training using two different procedures (see below), and reversed standard and target stimulus classes for the subjects.

All training procedures were modified from those of May et al. (1988; see also May 1987). We presented the first subject (m181) with a single target (a tonal scream) and a single standard (an arched scream). In contrast to May et al. (1988), we presented the target and standard at equivalent levels. Once m181 reliably responded to the target (responses $\geq 85\%$) and not the standard (responses $\leq 20\%$) for 3 consecutive days, we added a second stimulus pair to the test set. We added additional stimulus pairs once m181 reached criterion performance, until we presented eight targets and eight standards in each test session. We presented all stimuli monaurally, through the right earphone, based on evidence that adult rhesus show left hemisphere dominance for processing vocalizations (Hauser & Andersson 1994; Hauser et al. 1998). Because

the training duration for m181 was relatively long (53 days), we used a different protocol for m184.

For m184, signals could initially be distinguished based on intensity cues. In brief, we first trained m184 to report detection of a randomly ordered presentation of the eight scream targets (arched screams). Once she reliably reported arched scream detection, we added the standards (tonal screams) to the test set at a reduced level (-20 dB relative to targets). Once m184 responded to at least 80% of the targets, but less than 20% of the attenuated-level standards, we increased the standard level to -10 dB relative to the targets, and then to -5 dB relative to the targets. We then increased standard stimulus presentation level so that targets and standards were at equivalent levels. The total training time using these procedures was 30 days, thus, as predicted by May (1987), training appeared to progress more rapidly when all target and standard stimuli were introduced during the first test session. However, it is possible that individual differences were more important in determining the rate of task acquisition than the training paradigm.

For m184, we initially presented scream stimuli randomly to either the right or left ear. During initial classification testing (i.e. once target and standard scream stimuli were at identical levels), m184 responded significantly more rapidly ($P \leq 0.01$) and accurately ($P \leq 0.05$) when the test stimulus was presented to the right ear rather than the left. These data are consistent with previous observations of laterality in rhesus (Hauser & Andersson 1994; Hauser et al. 1998) and Japanese, *M. fuscata* (Petersen et al. 1984, 1978) macaques. Ear advantages were not evident beyond 2 months of testing. Therefore, we presented all stimuli monaurally through the right earphone during later testing.

Testing Procedures

We conducted test sessions 5 days per week for approximately 45 min daily. During trials, we presented standard scream stimuli at a rate of two/s for up to 9 s; we presented several different standards during each trial. During 75% of the trials, we presented a target following a random number of standards (i.e. 'test trials'). We provided reinforcement for correct release responses for 85% of these responses. During the remaining 25% of the trials, we monitored responses to standards (catch trials). We punished responses during catch trials with a brief (5 s) time-out.

We used hit rates and catch trial response rates to determine d' values for each test session (Elliott 1964). We selected a minimum d' value of 1.68, based on the d' associated with a hit rate of 80% and catch trial rate of 20%, as the minimum criterion for acceptable classification performance. When d' was at least 1.68 for three of the last four test sessions, and at least 75 food pellets were earned per session, we considered classification behaviour stable, and presented a generalization test (i.e. a 'probe session').

During probe sessions, we replaced one target in the daily test set with a novel stimulus (i.e. a 'probe' stimulus). We presented this stimulus on approximately 15% of the test trials. Responses to probe stimuli were not

reinforced or punished, and failure to release the response cylinder was not punished. All responses to the seven remaining targets were reinforced to maintain overall reinforcement at 85%. We conducted probe sessions no more than once weekly, and only if the animal had been tested the previous day. Following probe sessions, we classified stimuli as perceptual 'targets' when responses were greater than or equal to 70%. If responses were less than or equal to 30%, we classified the stimulus as a perceptual 'standard'. We labelled response rates between 30 and 70% perceptually 'ambiguous'.

We conducted probe sessions in cycles. We randomly determined the order of novel stimulus sessions within the first probe cycle for each subject, and reversed the second cycle. When a third cycle was tested, we again randomized probe order for each subject, and reversed the fourth cycle.

Apparatus

We transported the subjects to the laboratory using a chair restraint system (see Moody et al. 1970) and placed them inside a double-walled soundproof room (Industrial Acoustics Company, Bronx, New York, U.S.A.). We placed Telephonics (TDH 49P) earphones over each ear. The frequency response of the earphones was flat across the signal range that was tested. Above 6.25 kHz, the frequency response gradually declined. We moderately restricted motion of the subject's head to maintain earphone placement. The response cylinder was at hand level within easy reaching distance of the subject.

We used an Intel-based computer to control stimulus presentation and collect response data. We generated digitized scream vocalizations using a 12-bit digital-to-analogue converter set at a 25-kHz clock rate, then low-pass filtered the signals at 6 kHz. This filtering attenuated, but did not eliminate, energy above 6 kHz. In fact, when subjects were more recently tested with probe stimuli containing energy only above 6 kHz, these stimuli were detected, and discretely classified (unpublished observations). The filtered screams are depicted in all spectrographic illustrations. The screams were attenuated to a set signal level, amplified to match headphone impedance, and attenuated again to reduce any amplification noise.

Root-mean-square amplitude (rms: the square root of multiple instantaneous amplitude values squared then averaged) of each signal ranged from 1.2 to 2.0 V (see Table 1). Presentation level was set such that stimulus A4 was presented at approximately 80-dB sound pressure level (SPL). Because stimulus A4 was digitized at a lower rms value (1.31), actual stimulus levels ranged from approximately 80–83 dB SPL. A continuous masking noise (100–2000 Hz) was presented at 55 dB SPL. This was required as some signals contained low-frequency environmental noise. The scream vocalizations were clearly audible over the masker.

Stimuli

The scream vocalizations used during daily testing were arched (A1–A8, see Fig. 1) and tonal (T1–T8, see Fig. 2)

Table 1. Scream nomenclature

Scream*	RMS†	Vocalizing animal	
		Original label‡	Identification
A1	2.0	as2n8825.389	2n8
A2	2.0	as148032.052	148
A3	2.0	as1d1231.218	1d1
A4	1.31	as2d1232.613	2d1
A5	1.2	as1o0513.586	1o0
A6	2.0	as1l5847.154	1l5
A7	1.95	as1d1232.613	1d1
A8	1.8	as2o0513.586	2o0
T1	2.0	ts3n8825.381	3n8
T2	2.0	ts148032.320	148
T3	2.0	as3g6513.599§	3g6
T4	2.0	ts1g6513.599	1g6
T5	2.0	ts2n8825.381	2n8
T6	2.0	ts2h8301.612	2h8
T7	2.0	ts1n8825.389	1n8
T8	1.7	ts1g6532.040	1g6

*Scream labels were arbitrarily assigned to vocalizations.

†RMS (root mean square) amplitude provides a measure of overall amplitude across the entire duration of a signal.

‡Original labels are provided for comparison with reports from other investigators using these signals.

§This scream was initially labelled an arched scream. Spectrographic analysis resulted in reclassification as a tonal scream.

screams produced by free-ranging rhesus macaques on Cayo Santiago (see Table 1). Classification of screams was based on the criteria described by Gouzoules et al. (1984). In brief, arched screams were defined by initial sharply ascending frequency modulation, followed by descending frequency modulation, and calls were approximately 300 ms in duration or less. Tonal screams were identified by a longer segment with relatively little frequency modulation, and were approximately 300–500 ms in duration. They sometimes contained a frequency-modulated segment. Acoustic properties of the screams are summarized in Table 2. Synthetic screams were included in training and test sets in experiments 1 and 3. Probe stimuli included additional field-recorded screams (experiment 2) and synthetic stimuli (experiments 1 and 3).

Signal similarity has been compared using digital signal processing techniques, including spectrogram cross-correlation (e.g. Clark et al. 1987; Chabot 1988; Owren et al. 1992; Heaton et al. 1995; Mellinger & Clark 2000) and discriminant function analysis (e.g. Gouzoules & Gouzoules 1989; Hauser 1991; Rendall et al. 1998). Use of such techniques provides additional strength to signal classification schemes in that highly correlated signals are more likely to belong to the same class than are signals with low correlation values. We therefore attempted to analyse the field-recorded signals for spectrographic similarity using the automated pairwise correlation (CORMAT) feature of the Signal software package (Engineering Design, Belmont, Massachusetts, U.S.A.). Spectrogram similarity estimates provided by the CORMAT analysis (i.e. the correlation matrix) were

significantly disrupted by environmental noise in the recordings, however.

EXPERIMENT 1

Synthetic versions of primate vocal signals have been used in a number of investigations examining Japanese macaque perception of species-typical coo calls (e.g. May et al. 1988, 1989; Hopp et al. 1992; Le Prell & Moody 1997, 2000) as well as vervet monkey, *Cercopithecus aethiops*, perception of alarm calls (Owren 1990b). In these investigations, synthetic and field-recorded vocal signals were treated equivalently. Recent experiments using a habituation–discrimination procedure revealed rhesus monkeys treat synthetic vocalizations appropriately, and this includes the tonal scream (Hauser 1999). Thus, rhesus monkeys should classify synthetic vocal signals as similar to field-recorded signals in the laboratory.

In experiment 1, we used synthetic vocalizations to explore the perceptual categorization of screams by rhesus monkeys, explicitly testing the proposal that arched and tonal screams form discrete classes (Gouzoules et al. 1984). To do so, we trained animals to perform a behavioural response indicating their classification of scream signals. As long as the investigator-defined categories used to train the subjects were defined by the same criteria as the monkeys' natural categories, subjects' responses will reflect natural perceptual processes.

There are several potential experimental outcomes. First, if the animals naturally classify screams using discrete categories and the investigators select training stimuli that reflect these natural categories, the animals' responses should fall into discrete categories. Specifically, a scream treated as an arched scream by one animal should not be treated as a tonal scream by the second animal, and vice versa. Second, the animals might naturally respond to the calls along a graded continuum. If there is significant grading, subjects may extract a set of 'critical features' from the target stimuli presented during training. The subjects might then use these features to form one trained category consisting of a reinforced class of sounds. Alternatively, subjects may form two loosely defined categories of sounds based on multiple features that tend to define the two proposed scream classes. If so, investigator selection of reinforcement contingencies presumably will influence the subject's categorization scheme as the investigator will impose an artificial boundary defining discrete stimulus classes. Third and last, the data may fail to reflect categories of either a natural or a trained nature. Instead, the subjects may learn to respond based solely on reinforcement contingencies (i.e. memorize the specific stimuli for which responses are reinforced). If so, the subjects would not be expected to respond to any of the probe stimuli. This result would be surprising, given that another macaque species (*M. fuscata*) readily learned to categorize communicative stimuli (see Petersen et al. 1978, 1984; Beecher et al. 1979; May et al. 1988, 1989).

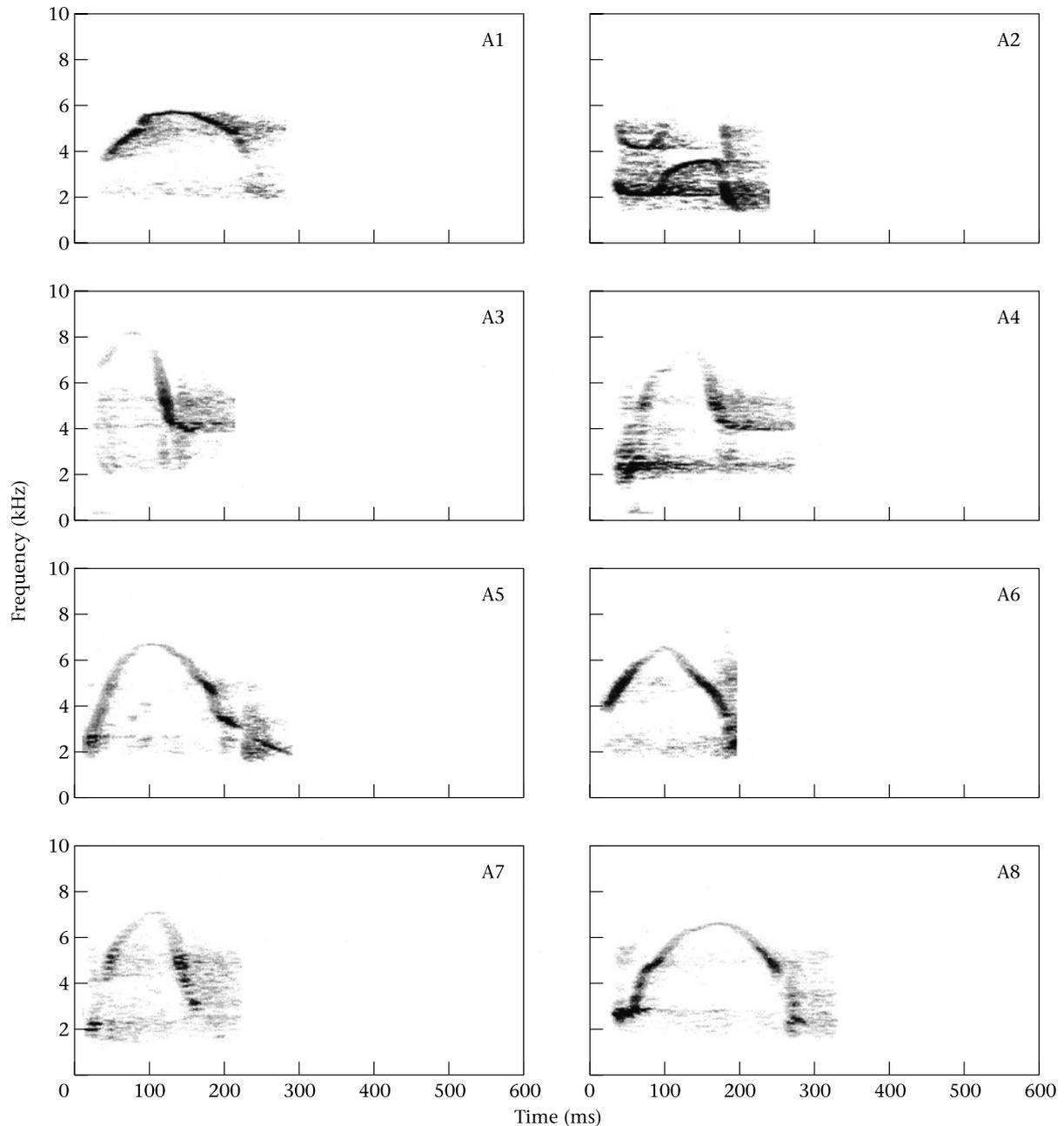


Figure 1. Spectrograms of rhesus monkey arched screams (A1–A8) recorded on Cayo Santiago. Spectrograms were produced by digitally reacquiring the filtered signals using the Daqarta signal analysis program (under development). Display resolution was set to depict a 55-dB range of signal intensities. The highest intensity signal components are depicted as the darkest areas within the spectrograms.

Methods

Procedure

We conducted daily testing as described under General Methods. Probe cycles included field recorded stimuli (cycle 1), synthetic ‘replica’ stimuli (cycles 2–5) and acoustically manipulated synthetic variants (cycle 6).

Stimuli

In addition to the field-recorded screams (A1–A8, T1–T8), we used five synthetic ‘replica’ stimuli modelled

after these screams (Synthetic A2, A5, T1, T6 and T7; see Fig. 3). We also presented variants of these synthetic stimuli (A2M, A5M1, A5M2, T1M1, T1M2, T6M and T7M; not depicted).

Synthetic replicas of the screams used in experiment 1 were computer generated using frequency and amplitude values for the fundamental frequency of phonation (F_0), and, when applicable, the upper harmonic (H2), derived from multiple Fast Fourier Transforms (FFTs) of the natural stimuli. Signal software was used to generate frequency and amplitude characteristics of each

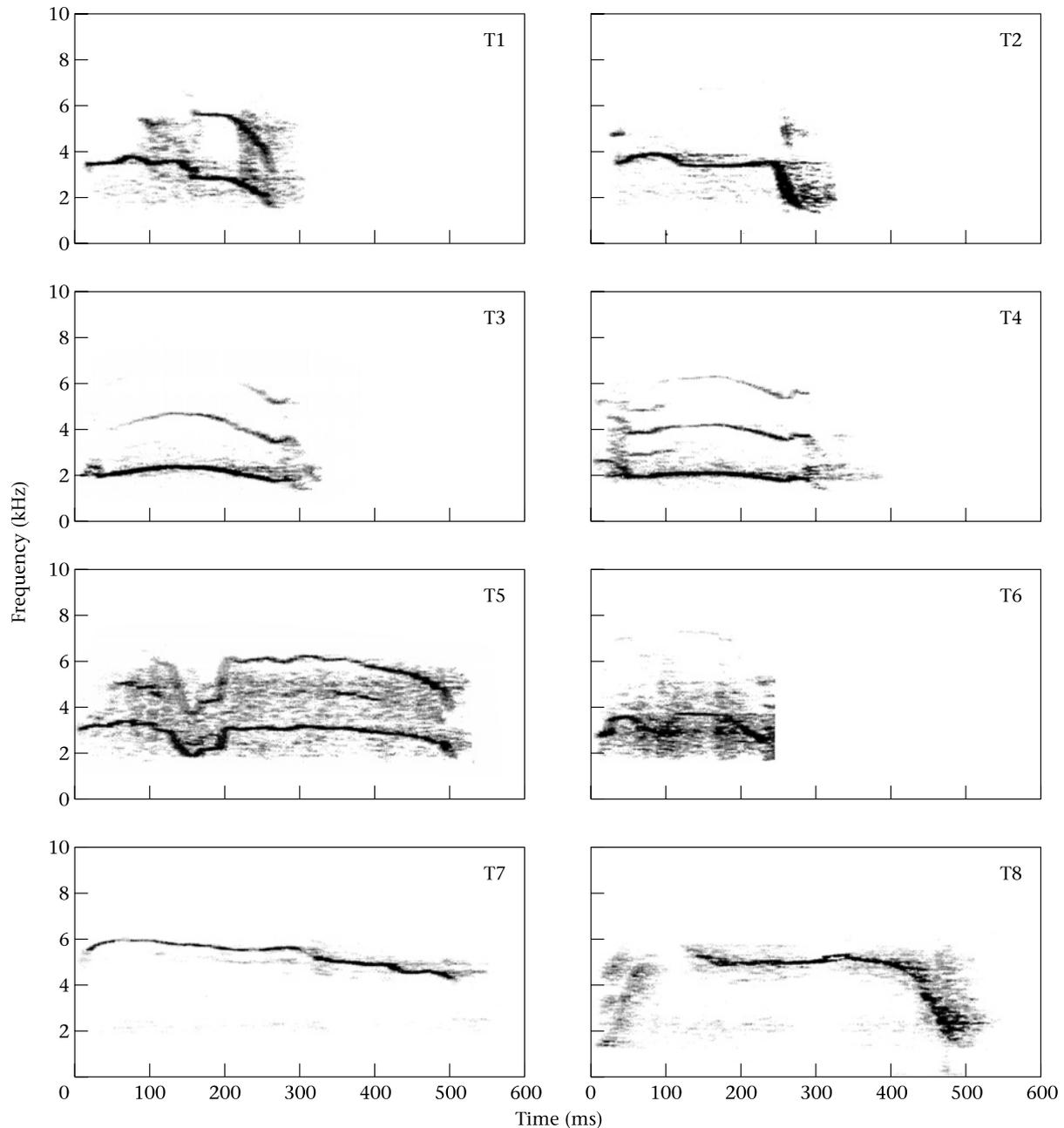


Figure 2. Spectrograms of rhesus monkey tonal screams (T1–T8) recorded on Cayo Santiago. Spectrograms were produced by digitally reacquiring the filtered signals using the Daqarta signal analysis program (under development). Display resolution was set to depict a 55-dB range of signal intensities. The highest intensity signal components are depicted as the darkest areas within the spectrograms.

harmonic component individually, and to combine F_0 and H2 to yield a single scream stimulus. In order to more closely approximate the field-recorded signals, each signal component was ‘jittered’ by randomly varying the frequency within a 10-Hz range. No attempt was made to model field-recorded environmental noise, as the masker hid environmental sounds.

Two human listeners judged the perceptual similarity of original and synthetic stimuli. The stimuli selected for probe presentation (i.e. synthetic versions of A2, A5, T1, T6 and T7) were those perceived as most similar to the field-recorded calls. Spectrograms for the synthetic

and field-recorded stimuli showed subtle differences. However, CORMAT-derived spectrographic correlation of the field-recorded signals and synthetic models were extremely high. Across the 15 synthetic models we generated, correlation values ranged from 0.63 to 0.97. For the five calls selected for testing, signal correlation values ranged from 0.86 to 0.97.

We conducted subtle acoustic manipulations on each synthetic replica (see Table 3). The purpose of the manipulations was to generate acoustically novel stimuli for which appropriate categorizations were known. Manipulations were accomplished using Signal software

Table 2. Acoustic properties of scream signals

Scream	Onset frequency* (Hz)	Maximum frequency* (Hz)	Ascending modulation†	Duration* (ms)
A1	3745	5524	2049	206.8
A2	2469	3552	1083	207.9
A3	6530	8000	1470	195.2
A4	2044	7265	5221	243.0
A5	2701	6878	4177	280.0
A6	4171	6723	2552	181.6
A7	1928	7110	5182	146.8
A8	2701	6646	3945	299.4
Average‡	3286	6462	3210	220.1
T1	3397	3707	310	291.6
T2	3552	3939	387	301.4
T3	1966	2353	387	319.5
T4	2044	2160	116	390.2
T5	3049	3359	310	525.9
T6	2701	3668	967	240.2
T7	5138	6066	928	552.9
T8	1541	5640	4099	547.3
Average§	2923	3861	938	396.1

*Acoustic parameters measured using cursor mode of Signal software package. All measurements refer to fundamental frequency of signal.

†Ascending frequency modulation defined as frequency range between onset frequency and maximum frequency (i.e. maximum frequency minus onset frequency).

‡Average measurement for arched scream vocalizations.

§Average measurement for tonal scream vocalizations.

to display the original signal on a computer monitor and using a computer drawing pad (Kurta XGT) and pen (XGT, cordless) to alter the frequency–time spectrum of the signal. Signal software was used to resynthesize modified stimuli based on the altered spectrograms.

Results

Following training, both subjects readily responded to a single scream class. Neither subject demonstrated high response levels to stimuli belonging to the alternative category. [Figure 4](#) depicts mean response percentages (\pm SD) to each vocalization during the five test sessions immediately prior to experiment 1. Responses to synthetic stimuli, and the field-recorded stimuli they were modelled after, are presented in [Fig. 5](#) (a: m181; b: m184). Responses to field-recorded stimuli were not altered as a function of probe contingencies (see [Fig. 4](#) for response data under standard test conditions).

Both m181 and m184 responded to replicas of target stimuli as if they were members of the target class. The percentage of responses to replicas of the standard (non-target) stimuli were variable and responses were elevated compared with response rates for the matching field-recorded stimuli. Indeed, standard class stimuli were often perceptually ‘ambiguous’. Although responses were elevated for standard synthetic vocalizations, the responses of both subjects differed between synthetic target and synthetic standard vocalizations. Responses to the synthetic variants, also summarized in [Fig. 5](#), were similar to responses to the synthetic replica stimuli.

Discussion

Although field studies reveal that rhesus monkeys do not differentially respond to noisy screams produced by kin and nonkin ([Rendall et al. 1998](#)), it seems that rhesus monkeys can distinguish the identity of animals producing coo calls (see [Hauser 1991](#); [Rendall et al. 1996](#)). We therefore included arched and tonal screams produced by the same animal to guarantee that subjects could not learn to attend only to the identity of the vocalizer. For example, the same animal produced arched scream A1 and tonal scream T5 (see [Table 1](#)). A second animal produced calls A2 and T2. Both subjects classified different signals produced by the same individual into different acoustic classes.

The strongest evidence for discrete perceptual categories of arched and tonal scream signals would be provided by symmetric responses to the novel synthetic stimuli. That is, if one animal consistently identifies arched screams as targets, the other animal should classify these signals as nontargets, while identifying tonal screams as targets. While responses to the training stimuli were clearly symmetric, responses to synthetic stimuli were less so. Response rates for standard-class synthetic stimuli were somewhat elevated for both animals. Thus, it appears that each subject formed a target category in which certain stimuli were readily identified as belonging to the target class. In contrast, the standard class was less rigorously defined and stimuli belonging to this class were often treated ambiguously. It is unlikely that the animals failed to learn any categorization scheme, responding on the basis of memorized (and reinforced)

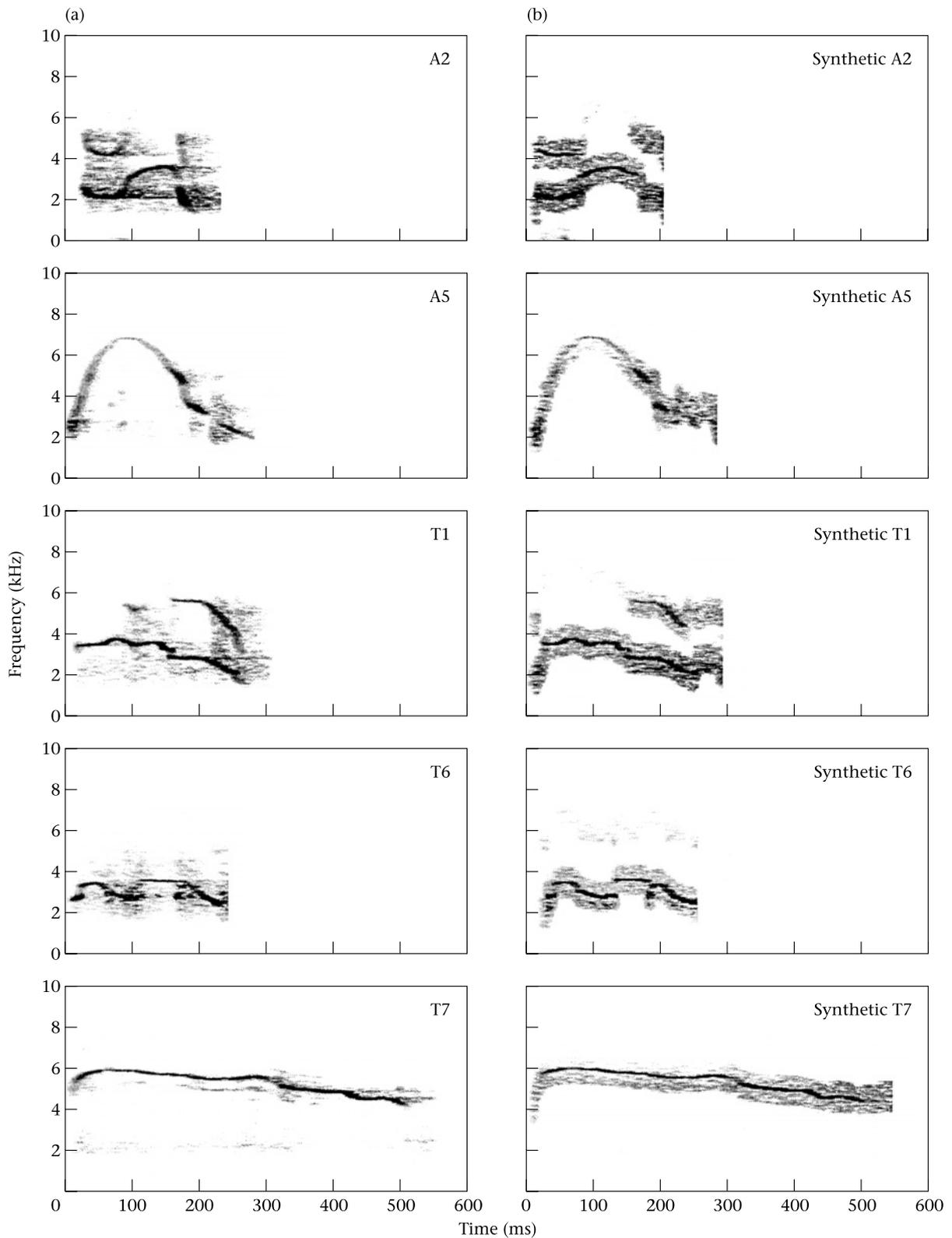


Figure 3. Spectrograms of field-recorded arched (A2, A5) and tonal (T1, T6, T7) scream vocalizations (a) and synthetic models ('replicas') of those screams (b). Spectrograms were produced by digitally reacquiring the filtered signals using the Daqarta signal analysis program (under development). Display resolution was set to depict a 55-dB range of signal intensities. The highest intensity signal components are depicted as the darkest areas within the spectrograms.

Table 3. Synthetic scream manipulations

Synthetic scream	Modified scream*	Manipulation
A2	A2M	Smoothed harmonics
A5	A5M1	Upper frequency values decreased by 1000 Hz†
A5	A5M2	Upper frequency values increased by 1000 Hz†
T1	T1M1	Smoothed harmonics to lower frequency values‡
T1	T1M2	Smoothed harmonics to upper frequency values‡
T6	T6M	Smoothed harmonics
T7	T7M	All frequency values decreased by 1000 Hz§

*Scream labels were assigned by adding an 'M' to the original label to indicate that the signal was manipulated. For screams manipulated more than once, a single digit following the 'M' distinguishes manipulations.

†Altering the upper frequency values changed the frequency bandwidth as well as the rate of frequency change.

‡T1 was characterized by a distinct drop in frequency values midway through the call. The lower harmonic was smoothed using either the lower frequency values during the first half of the call (T1M1) or the higher frequency values during the latter half of the call (T1M2). Exact harmonic multiples were added to the smoothed components. Components of the T1M1 and T1M2 stimulus differed by approximately 500 Hz.

§Frequency values were decreased by approximately 1000 Hz across call duration. This manipulation did not alter frequency bandwidth or rate of frequency change.

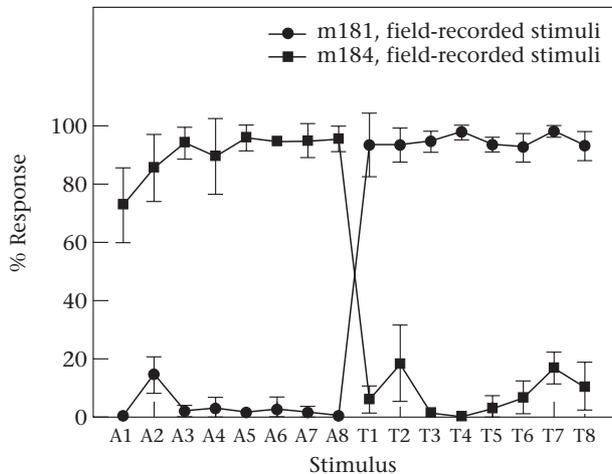


Figure 4. Subjects were trained to report the detection of either arched (m184) or tonal (m181) scream vocalizations. Mean (\pm SD) percentage of responses to all training stimuli (A1–A8, T1–T8) during the 5-day period immediately preceding the onset of experiment 1 are presented.

exemplars alone, given that the synthetic stimuli differed from the field-recorded stimuli. There were spectrographic differences between the synthetic and field-recorded stimuli (see Fig. 3 for comparison), and human listeners could generally discriminate field-recorded screams from synthetic replicas. Moreover, responses to manipulated variants of the synthetic stimuli were high for target stimuli. Taken together, these results do not support the conclusion that arched and tonal screams form perceptually discrete categories.

EXPERIMENT 2

To explore further whether subjects could form discrete arched and tonal scream classes, we measured the subjects' responses to novel (i.e. unfamiliar) field-recorded

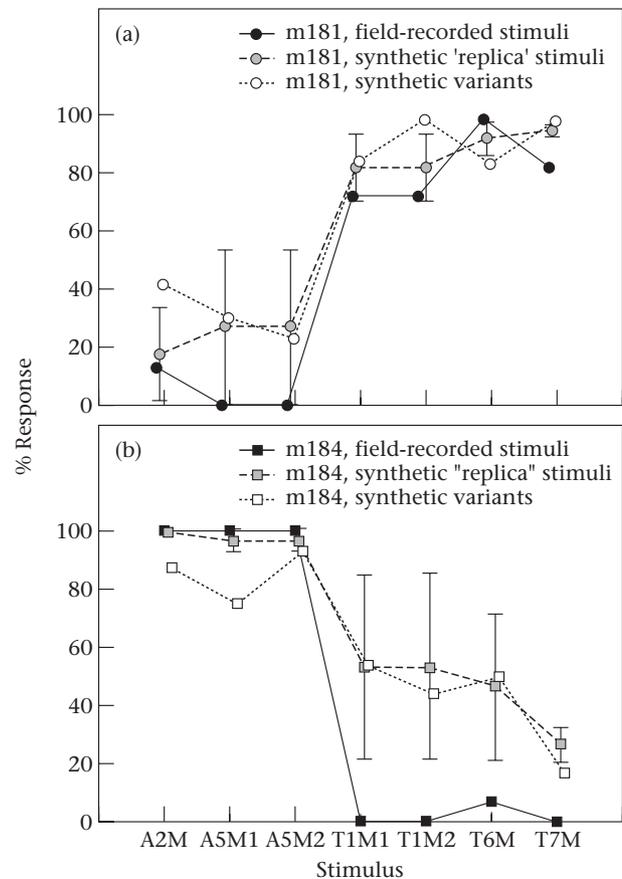


Figure 5. Mean percentage of responses to field-recorded scream vocalizations and synthetic 'replica' stimuli modelled after the original stimuli (a: m181; b: m184). Each field-recorded scream was tested during one probe session, and each synthetic stimulus was tested during four probe sessions. Error bars indicate standard deviation of responses to synthetic calls. Each manipulated stimulus (A2M, A5M1, A5M2, T1M1, T1M2, T6M, T7M) was presented during only one test session. Responses to field-recorded and synthetic replica stimuli are plotted with the corresponding manipulated stimulus.

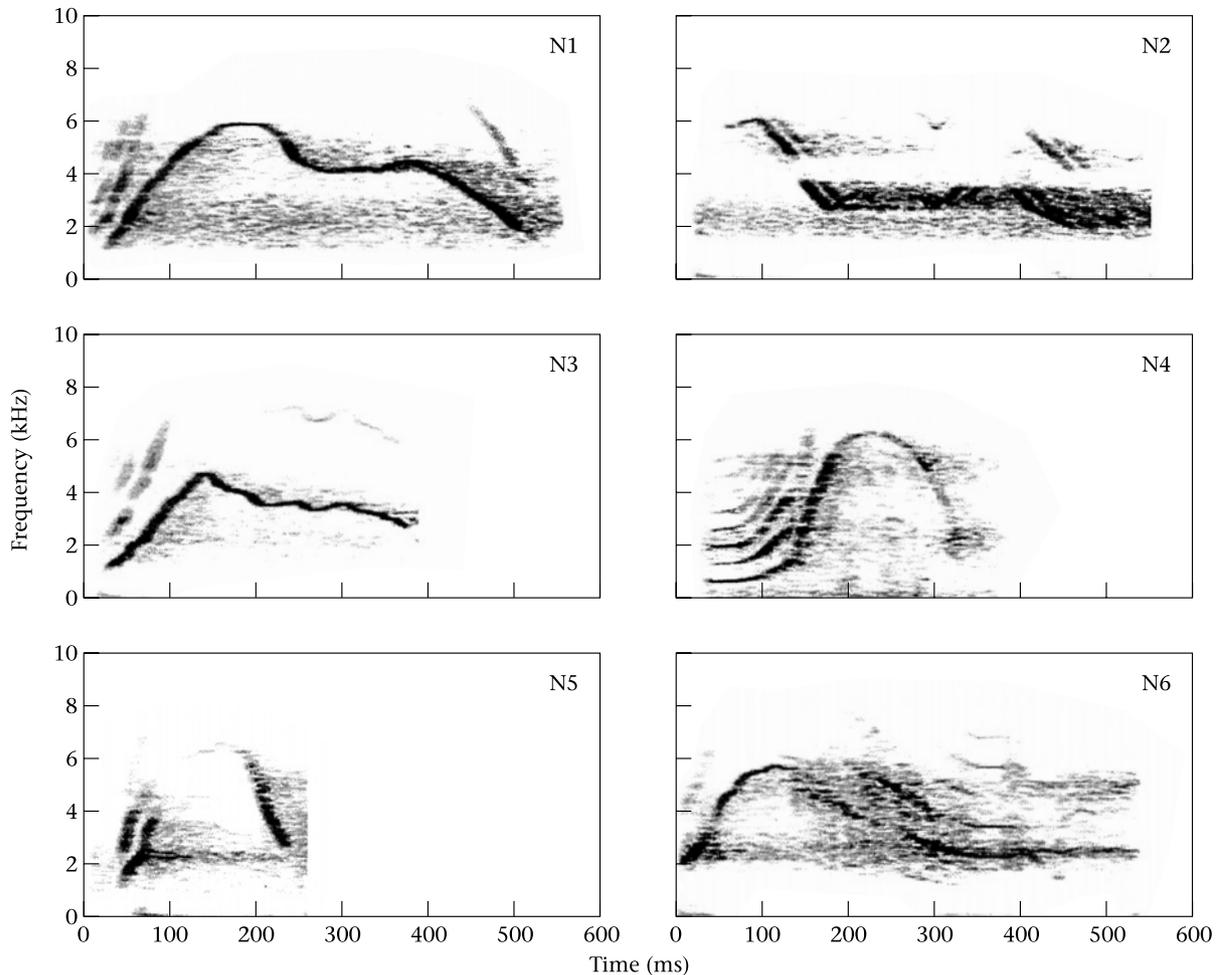


Figure 6. Spectrograms of unfamiliar (novel) scream vocalizations (N1–N6) used during generalization testing. Screams were recorded on Cayo Santiago. Spectrograms were produced by digitally reacquiring the filtered signals using the Daqarta signal analysis program. Display resolution was set to depict a 55-dB range of signal intensities. The highest intensity signal components are depicted as the darkest areas within the spectrograms.

screams during generalization tests. Discrete classification of these signals would provide evidence that specific vocal signal categories were learned (for discussion, see Prosen et al. 1990; see also May et al. 1988). Failure to demonstrate that subjects classify novel exemplars into discrete scream categories would support the hypothesis that screams represent a graded class of vocalizations.

To exclude further the possibility that subjects were responding based on caller identity, we included vocalizations produced by individuals whose calls were used in the training set. Specifically, the probe set included novel stimulus N2, produced by the same animal responsible for A3 and A7, and novel stimulus N4, produced by the same animal responsible for T1 (see Tables 1 and 4). If caller identity is an important characteristic learned by these subjects, then stimulus N2 should be treated similarly to A3 and A7 (i.e. as an arched scream) and stimulus N4 should be treated similarly to T1 (i.e. as a tonal scream). In contrast, if caller identity is perceived as irrelevant in this task, then N2 will be classified as a tonal scream, while N4 will be classified as an arched scream, based on the criteria of Gouzoules et al. (1984).

Method

Stimuli

We presented six field-recorded screams (N1–N6, see Fig. 6) during probe sessions. N1–N3 were generally characteristic of tonal screams as they contained relatively long tonal segments, while N4–N6 were more characteristic of arched screams. However, the majority of these screams were clearly unlike the training stimuli in that they contained both tonal-like and arched-like components. Although atypical of the categories described by Gouzoules et al. (1984), these screams were natural calls produced by free-ranging animals. In addition, these screams had the best signal-to-noise ratio of the signals available to us at that time. Previous nomenclature used to identify the N1–N6 signals is presented in Table 4.

Procedure

We conducted daily testing as described under General Methods. We conducted probe sessions until subjects had been tested with each novel stimulus during four separate probe sessions. Due to an error in setting up the test

Table 4. Novel scream stimuli

Scream*	Vocalizing animal		
	RMS†	Original label‡	Identification
N1	2.0	ts1j7725.283	1j7
N2	2.0	ts1d1232.687	1d1
N3	2.0	ts1a6703.231	1a6
N4	1.9	as3n8825.389	3n8
N5	2.0	as3d1232.614	3d1
N6	2.0	as1h787a.556	1h7

*Scream labels were arbitrarily assigned to vocalizations.

†RMS (root mean square) provides a measure of overall signal amplitude.

‡Original labels are provided for comparison with reports from other investigators using these signals.

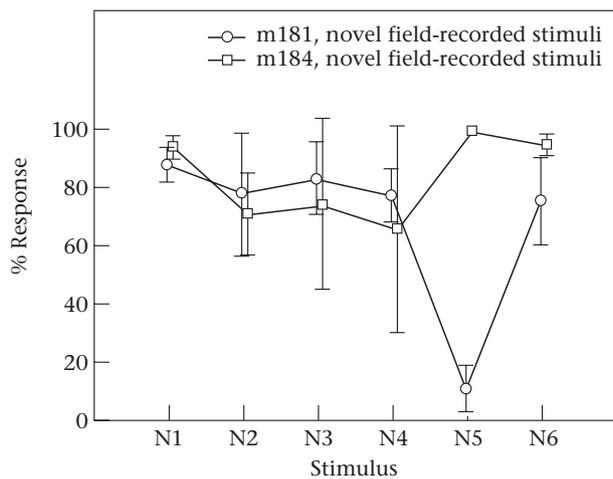


Figure 7. Mean (\pm SD) percentage of responses to unfamiliar scream vocalizations (N1–N6). Each stimulus was presented during four separate probe sessions.

session parameters, we treated probe stimuli presented to m181 in probe cycles 1 and 2 as standard stimuli; that is, we punished responses with a brief time-out.

Results

Mean percentage of responses (\pm SD) to each of the unfamiliar vocalizations are presented in Fig. 7. Stimulus N5 was classified as an arched scream by both m184 (99% arched responses) and m181 (18% tonal responses). In contrast, responses to the other probe stimuli did not differ as a function of training contingencies (i.e. which stimulus class served as the target class).

The responses of m181 to probe stimuli were generally quite consistent across probe cycles. The greatest variation was seen for stimulus N2. Some of this variability can presumably be attributed to the punishment of probe stimulus responses in cycles 1 and 2. Responses to stimulus N2 were 76% (cycle 1) and 50% (cycle 2) when responses to this stimulus were punished, and 86% (cycle 3) and 100% (cycle 4) when probe parameters were correctly applied. Responses of m184 were stable across

probe sessions, except for the N3 and N4 stimuli. The large response variation of m184 to stimuli N3 and N4 represents a systematic shift in response rates. Whereas responses to N3 and N4 ranged from 85 to 100% during the first two probe cycles, responses ranged from 21 to 54% during the latter cycles. Systematic shifts in responses to other probe stimuli were not evident.

Discussion

There was no evidence that responses were based on previous experience with signals produced by the same caller. With the exception of the N5 stimulus, responses to probe stimuli did not reveal discrete classification. Specifically, both subjects generally responded to each probe stimulus as if it were a member of the target class. Thus, given the constraints imposed by our task, our data suggest that subjects were responding to these stimuli as a function of stimulus novelty rather than either trained or inherently meaningful communication categories.

Prior to generalization testing, we noted that the N1–N6 signals tended to contain both tonal-like and arched-like segments. Therefore, it is not necessarily surprising that N5 was the only stimulus classified differently by the two subjects. This call was the only exemplar to be unambiguously assigned to a single category either acoustically (see Fig. 6) or perceptually (see Fig. 7). We believe the relative ambiguity of our novel probes probably contributed to the subjects' failure to discretely classify the novel stimuli. That is, because the signals contained elements of each subjects' target class, both subjects may have responded based on the identification of critical features extracted from the training stimuli. The observation that such screams are produced in natural settings, combined with ambiguity in subjects' classificatory responses, provides additional evidence that arched and tonal screams do not form discrete categories. Experiment 3 therefore tests the hypothesis that subjects' responses are based on critical features extracted from the overlapping arched and tonal scream categories.

EXPERIMENT 3

Semifree-ranging (Hauser 1999) and captive (experiment 1) rhesus monkeys treat synthetic scream vocalizations similarly to field-recorded exemplars. We therefore manipulated synthetic screams to determine whether subjects' classificatory responses are mediated by highly salient acoustic features associated with training exemplars. We generated novel exemplars of arched and tonal screams using a stimulus-averaging procedure. By averaging acoustic parameters from multiple exemplars, we maintained the original features of the fundamental frequency (F_0) contour, but altered other acoustic variables. We use the term 'prototypes' to refer to these stimuli.

Method

Stimuli

Stimuli in the daily training sets included synthetic (A2, A5, T1, T6 and T7; see Fig. 3) and field-recorded (A1,

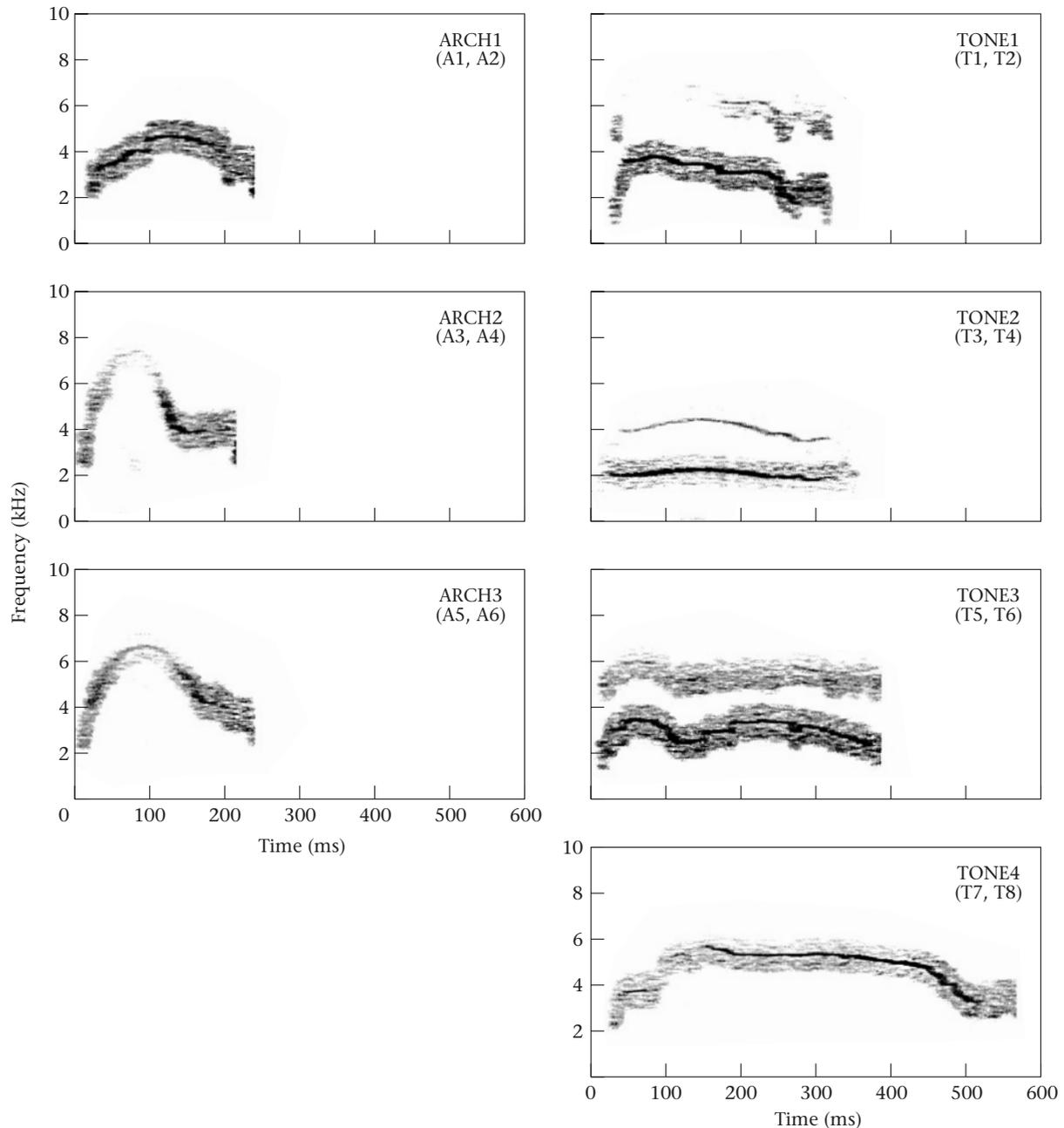


Figure 8. Spectrograms of 'prototype' screams (ARCH1, ARCH2, ARCH3, TONE1, TONE2, TONE3, TONE4) that were synthesized based on averaged acoustic features of field-recorded screams (A1–A6, T1–T8). For example, ARCH1 is based on averaging of features of A1 and A2. Spectrograms were produced by digitally reacquiring the filtered signals using the Daqarta signal analysis program. Display resolution was set to depict a 55-dB range of signal intensities. The highest intensity signal components are depicted as the darkest areas within the spectrograms.

A3, A4, A6, A7, T2, T3, T4 and T5; see [Figs 1 and 2](#)) stimuli. We tested seven synthetic prototypes (ARCH1, ARCH2, ARCH3, TONE1, TONE2, TONE3 and TONE4; refer to [Fig. 8](#)) using probe procedures.

We selected stimulus pairs to be averaged (i.e. 'source' stimuli) based on relative similarities in F_0 contour. First, 99 512-point Fast Fourier Transforms (FFTs) were carried out on the field-recorded stimuli. Across stimuli, spacing between successive FFTs ranged from 1.81 to 5.53 ms

(mean=3.15 ms). We then extracted frequency and amplitude values from these FFT analyses.

Within each pair of source stimuli, we first determined the mean duration of the field-recorded signals and then scaled each signal to this duration. Because we sampled an equal number of time points in the FFTs in each stimulus, we accomplished scaling of signal duration by assigning new time point values to the existing frequency and amplitude contours. For each new time point in the

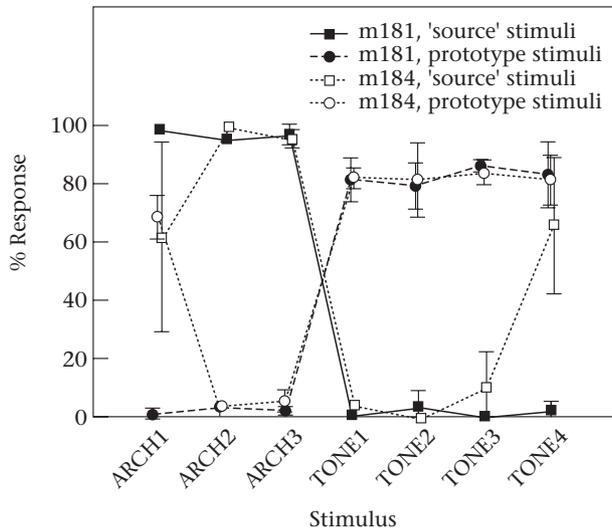


Figure 9. Mean (\pm SD) responses to presentations of prototype scream stimuli (ARCH1, ARCH2, ARCH3, TONE1, TONE2, TONE3, TONE4) during two probe test sessions. Percentage of responses to each 'source' stimulus were also measured. Source stimuli were those after which the prototype stimuli were modelled.

prototype stimulus, we then averaged the existing frequency and amplitude contours and used Signal software to synthesize the new signals. Each harmonic component was 'jittered' by randomly varying the frequency within a 1–13-Hz range to better approximate field-recorded species-typical scream vocalizations.

Procedure

Prior to generalization testing, we inserted synthetic replica stimuli into the training set in place of the field-recorded stimuli after which they were modelled. Thus, we reinforced responses to synthetic targets and punished responses to synthetic standards. Subject m181 treated all synthetic stimuli appropriately (responses to targets at least 90%, responses to nontarget stimuli less than 1%) during the first test session. Responses to target stimuli were lower for m184 on the first day of testing (approximately 60% responses to targets). Her responses increased to at least 80% by the second day of testing. We trained subjects with this stimulus set for approximately 3 weeks, and then initiated probe sessions.

We conducted two probe cycles containing prototype stimuli. During each probe session, we also presented and assessed responses to one of the two source stimuli.

Results

Percentage of responses to source and prototype stimuli are depicted in Fig. 9. All source stimuli were discretely classified as either arched or tonal. Five of the seven prototype stimuli were also classified as either tonal or arched. Neither subject treated the ARCH1 stimulus as distinctively arched or tonal (m181: 69% tonal responses; m184: 62% arched responses) however. Responses of m184 to ARCH1 (39–85%) and TONE4 (50–83%) were variable. In contrast, m181 classified TONE4 as tonal (82% tonal responses).

Discussion

The fact that subjects responded ambiguously to ARCH1 (m181, m184) and TONE4 (m184) raises further questions about the discrete classification of screams. Although the prototype screams were synthetically generated and may not perfectly represent naturally produced vocalizations, we obtained similarly ambiguous response patterns when we presented field-recorded screams in experiment 2, and both field and laboratory results suggest that rhesus treat synthetics as valid exemplars.

Examining the spectrograms in Fig. 8 reveals that ARCH1 was the poorest arched scream prototype as this scream had a much slower rate of frequency change than the other arched scream prototypes. Following the terminology of Gouzoules et al. (1984), this stimulus was defined by a shallower frequency rise and descent, or less frequency modulation. In addition, the maximum frequency of the F_0 contour (hereafter referred to as F_0 MAX) was much lower than for the other arched prototypes. Thus, the total frequency bandwidth was narrower as well. These results suggest that scream classification is mediated by at least one of several potentially critical features (i.e. rate of frequency change, F_0 MAX, and frequency bandwidth).

TONE4 was classified as a tonal scream by m181, presumably due to the long tonal component in this signal. Responses of m184 to TONE4, indicating arched scream detection, were variable, however. Although this stimulus did not appear to have a more rapid rate of frequency change or a wider frequency bandwidth than did other tonal prototypes, F_0 MAX was approximately 2 kHz higher than in any other tonal prototype. In addition, TONE4 was the only tonal prototype that did not contain a harmonic component. These results thus support the suggestion that F_0 MAX could be a critical classification feature, and further suggest that harmonic structure could be perceptually important for arched and tonal scream classifications.

GENERAL DISCUSSION

Both subjects learned to classify arched and tonal screams. To determine whether the subjects had learned abstractions of the arched and tonal categories described by Gouzoules et al. (1984), we presented subjects with novel synthetic screams in experiments 1 and 3, and novel field-recorded screams in experiment 2. If the arched and tonal scream classes are perceptually discrete, as suggested by Gouzoules et al. (1984), then subjects should classify the novel stimuli into discrete categories. Results show, however, that responses to the nontarget standards were elevated in experiment 1, only one stimulus was discretely classified in experiment 2, and some, but not all, stimuli were discretely classified in experiment 3. The response data thus suggest that subjects were responding to the standard stimuli as a function of a loosely defined category, an interpretation consistent with a graded communication class characterized by multiple salient acoustic features. These results force us to

question the hypothesis that screams are naturally divided into discrete classes.

Experiment 3 provided the strongest evidence that exemplars could be discretely classified into arched and tonal categories. Identification of two scream exemplars that subjects failed to categorize discretely led to the hypothesis that classification depends on multiple features such as harmonic structure, rate of frequency change and F_0 MAX. Some of these acoustic features may vary with signaller arousal, which is consistent with Gouzoules et al.'s (1984) observation that screams with varied acoustic morphology are produced as a function of the severity of aggression directed at the vocalizer. Parametrically manipulating acoustic features of these spectrally complex vocal signals may reveal the particular salience of individual acoustic features and allow closer examination of grading across categories. Such manipulations are difficult, however, because changing one acoustic feature often changes other features. For example, increasing the rate of frequency change results in higher F_0 MAX values and increases frequency bandwidth. Similarly, generating harmonic components alters the rate of frequency change. Isolating the salience of particular features represents a challenge for future research.

The failure of subjects to classify screams symmetrically into discrete classes and the variety of acoustic features that seem to influence scream classification in a laboratory setting appears to reflect the production and perception of a graded vocal category. Indeed, the availability of screams containing both tonal and arched features is consistent with the graded vocalization scheme proposed by Rowell (1962; see also Rowell & Hinde 1962). This finding stands in contrast to the observation that under field conditions, rhesus respond (latency and duration of orienting response) quite discretely to arched and tonal screams, and that acoustic analyses reveal a relatively low incidence of ambiguously classified calls (Gouzoules et al. 1984).

Given the reported differences in adult responses to different scream signals (Gouzoules et al. 1984), one might conclude that different scream types evolved to alert allies in situations requiring aid against an aggressor. The overlap in the contexts associated with scream production, however, makes a rigid classification of screams by putative meaning seem unlikely. Indeed, although our laboratory tests clearly differ from the perceptual tests conducted in the field, we must question the claim that rhesus perceive discrete scream categories given that our subjects developed only loosely defined acoustic categories. The final question, then, is whether the arched and tonal categories apparently learned by subjects in these experiments were acquired through training (see Stebbins 1990), or were a naturally graded class of vocalizations that training failed to alter. Presumably, multi-dimensional scaling (MDS) techniques, used to analyse perceived signal similarity in the absence of specific categorization training (Dooling 1986, 1989), could answer this question. This approach has been used to examine coo call perception in Japanese macaques (Le Prell & Moody 1997, 2000).

Acknowledgments

Portions of this research were presented at the 1998 Midwinter Meeting of the Association for Research in Otolaryngology (Le Prell & Moody 1998). This research was supported by NIH-NIDCD Program Project Grant P01-DC00078 and NIH-NIDCD Institutional Training Grant T32-DC00011 and a young investigators award from the National Science Foundation (M.D.H.). Preparation of this manuscript was additionally supported by National Research Service Award F32 DC00367. The authors gratefully acknowledge the contributions of Catherine Thompson, who had the primary responsibility for the training and daily testing of these animals, and Bob Masta, who designed and provided the Daqarta (data acquisition and real time analysis) program used to digitally acquire and depict the spectrograms. We also thank Rebecca Coste, Edwin Labut and Andrew Niemiec for periodic assistance. Finally, we thank Michael Owren for valuable comments on an earlier version of this manuscript. The research presented here was described in Animal Research Protocol No. 6041, approved on 16 November 1995 by the Institutional Animal Care and Use Committee of the University of Michigan.

References

- Beecher, M., Petersen, M., Zoloth, S., Moody, D. & Stebbins, W. 1979. Perception of conspecific vocalizations by Japanese macaques: evidence for selective attention and neural lateralization. *Brain, Behavior and Evolution*, **16**, 443–460.
- Brown, C. H., Beecher, M. D., Moody, D. B. & Stebbins, W. C. 1979. Locatability of vocal signals in Old World monkeys: design features for the communication of position. *Journal of Comparative and Physiological Psychology*, **93**, 806–819.
- Chabot, D. 1988. A quantitative technique to compare and classify humpback whale (*Megaptera novaengliae*) sounds. *Ethology*, **77**, 89–102.
- Cheney, D. L. & Seyfarth, R. M. 1988. Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, **36**, 477–486.
- Cheney, D. L. & Seyfarth, R. M. 1999. Recognition of other individuals' social relationships by female baboons. *Animal Behaviour*, **58**, 67–75.
- Cheney, D. L., Seyfarth, R. M. & Silk, J. M. 1995. The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Animal Behaviour*, **50**, 249–257.
- Clark, C. W., Marler, P. & Beeman, K. 1987. Quantitative analysis of animal vocal phonology: an application to swamp sparrow song. *Ethology*, **76**, 101–115.
- Dooling, R. J. 1986. Perception of vocal signals by budgerigars (*Melopsittacus undulatus*). *Experimental Biology*, **45**, 195–218.
- Dooling, R. J. 1989. Perception of complex, species-specific vocalizations by birds and humans. In: *The Comparative Psychology of Audition: Perceiving Complex Sounds* (Ed. by R. J. Dooling & S. H. Hulse), pp. 423–445. Hillsdale, New Jersey: L. Erlbaum.
- Elliott, P. B. 1964. Appendix 1: Tables of d. In: *Signal Detection and Recognition by Human Observers: Contemporary Readings* (Ed. by J. A. Swets), pp. 651–685. New York: J. Wiley.
- Fajzi, K., Reinhardt, V. & Smith, M. 1989. A review of environmental enrichment strategies for singly caged nonhuman primates. *Lab Animal*, **18**, 23–35.
- Fischer, J. 1998. Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*, **55**, 799–807.

- Fischer, J., Hammerschmidt, K. & Todt, D. 1998. Local variation in Barbary macaque shrill barks. *Animal Behaviour*, **56**, 623–629.
- Gouzoules, S. & Gouzoules, H. 1989. Design features and developmental modification of pigtail macaque, *Macaca nemestrine*, agonistic screams. *Animal Behaviour*, **37**, 383–401.
- Gouzoules, S., Gouzoules, H. & Marler, P. 1984. Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Animal Behaviour*, **32**, 182–193.
- Green, S. 1975. Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): a field study. In: *Primate Behavior*. Vol. 4 (Ed. by L. Rosenbloom), pp. 1–102. New York: Academic Press.
- 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.
- Hauser, M. D. 1999. The evolution of a lopsided brain: asymmetries underlying facial and vocal expressions in nonhuman primates. In: *The Design of Animal Communication* (Ed. by M. D. Hauser & M. Konishi), pp. 597–628. Cambridge, Massachusetts: MIT Press.
- Hauser, M. D. 1998. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Animal Behaviour*, **55**, 1647–1658.
- Hauser, M. D. & Andersson, K. 1994. Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkey: field experiments. *Proceedings of the National Academy of Sciences U.S.A.*, **91**, 3946–3948.
- Hauser, M. D. & Marler, P. 1993. Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behavioral Ecology*, **4**, 194–205.
- Hauser, M. D., Agnetta, B. & Perez, C. 1998. Orienting asymmetries in rhesus monkeys: the effect of time-domain changes on acoustic perception. *Animal Behaviour*, **56**, 41–47.
- Heaton, J. T., Farabaugh, S. M. & Brauth, S. E. 1995. Effect of syringeal denervation in the budgerigar (*Melopsittacus undulatus*): the role of the syrinx in call production. *Neurobiology of Learning and Memory*, **64**, 68–82.
- Hopp, S. L., Sinnott, J. M., Owren, M. J. & Petersen, M. R. 1992. Differential sensitivity of Japanese macaques (*Macaca fuscata*) and humans (*Homo sapiens*) to peak position along a synthetic coo call continuum. *Journal of Comparative Psychology*, **106**, 128–136.
- Kaplan, J. R. 1977. Patterns of fight interference in free-ranging rhesus monkeys. *American Journal of Physical Anthropology*, **47**, 279–288.
- Kaplan, J. R. 1978. Fight interference and altruism in rhesus monkeys. *American Journal of Physical Anthropology*, **49**, 241–250.
- Kitko, R., Gesser, D. & Owren, M. J. 1999. Noisy screams of macaques may function to annoy conspecifics. *Journal of the Acoustical Society of America*, **106**, 2221.
- 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. 1998. Discrimination of natural and synthetic scream call vocalizations by adult female rhesus macaques. In: *Association for Research in Otolaryngology, 21st Midwinter Meeting Abstracts* (Ed. by G. R. Popelka), page 102. Des Moines, Iowa: Association for Research in Otolaryngology.
- Le Prell, C. G. & Moody, D. B. 2000. Factors influencing the salience of temporal cues in the discrimination of Japanese monkey coo calls. *Journal of Experimental Psychology: Animal Behavior Processes*, **26**, 261–273.
- May, B. 1987. Significant features and perceptual categories in the vocal communication of Japanese macaques. Ph.D. thesis, University of Michigan.
- May, B., Moody, D. & Stebbins, W. 1988. The significant features of Japanese macaque coo sounds: a psychophysical study. *Animal Behaviour*, **36**, 1432–1444.
- May, B., Moody, D. & Stebbins, W. 1989. Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *Journal of the Acoustical Society of America*, **85**, 837–847.
- Mellinger, D. K. & Clark, C. W. 2000. Recognizing transient low-frequency whale sounds by spectrogram correlation. *Journal of the Acoustical Society of America*, **107**, 3518–3529.
- Moody, D. B., Stebbins, W. C. & Miller, J. M. 1970. A primate restraint and handling system for auditory research. *Behavioral Research Methods and Instrumentation*, **2**, 180–182.
- National Research Council, Institute of Laboratory Animal Resources 1996. *Guide for the Care and Use of Laboratory Animals*. Washington, D.C.: National Academy Press.
- Owren, M. J. 1990a. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: I. Natural calls. *Journal of Comparative Psychology*, **104**, 20–28.
- Owren, M. 1990b. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): II. Synthetic calls. *Journal of Comparative Psychology*, **104**, 29–40.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M. & Cheney, D. L. 1992. Food calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally raised offspring, and offspring cross-fostered between species. *Behaviour*, **120**, 218–231.
- Petersen, M. R., Beecher, M. D., Zoloth, S., Green, S., Moody, D. B. & Stebbins, W. C. 1978. Neural lateralization of species-specific vocalizations by Japanese macaques (*Macaca fuscata*). *Science*, **202**, 324–327.
- Petersen, M. R., Beecher, M. D., Zoloth, S., Green, S., Marler, P., Moody, D. B. & Stebbins, W. C. 1984. Neural lateralization of vocalizations by Japanese macaques: communicative significance is more important than acoustic structure. *Behavioral Neuroscience*, **98**, 779–790.
- Prosen, C. A., Moody, D. B. & Stebbins, W. C. 1990. A changing focus of animal perception: from abilities to proclivities. *International Journal of Comparative Psychology*, **4**, 5–23.
- Rendall, D., Rodman, P. S. & Emond, R. E. 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, **51**, 1007–1015.
- Rendall, D., Owren, M. J. & Rodman, P. S. 1998. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *Journal of the Acoustical Society of America*, **103**, 602–614.
- Rendall, D., Seyfarth, R. M., Cheney, D. L. & Owren, M. J. 1999. The meaning and function of grunt variants in baboons. *Animal Behaviour*, **57**, 583–592.
- Rowell, T. E. 1962. Agonistic noises of the rhesus monkey (*Macaca mulatta*). *Symposium of the Zoological Society of London*, **8**, 91–96.
- Rowell, T. E. & Hinde, R. A. 1962. Vocal communication by the rhesus monkey (*Macaca mulatta*). *Proceedings of the Zoological Society of London*, **138**, 279–294.
- Schapiro, S. J., Brent, L., Bloomsmith, M. A. & Satterfield, W. C. 1991. Enrichment devices for nonhuman primates. *Lab Animal*, **20**, 22–28.
- Seyfarth, R. M., Cheney, D. L. & Marler, P. 1980a. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, **210**, 801–803.
- Seyfarth, R. M., Cheney, D. L. & Marler, P. 1980b. Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Animal Behaviour*, **28**, 1070–1094.
- Stebbins, W. C. 1990. Uncertainty in the study of comparative perception: a methodological challenge. In: *BioMethods: Vol 6. Methods in Comparative Psychoacoustics* (Ed. by H. P. Saluz, T. Meier, G. M. Klump, R. J. Dooling, R. R. Fay & W. C. Stebbins), pp. 331–341. Boston: Birkhauser Verlag.

- Struhsaker, T. T.** 1967. Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In: *Social Communication Among Primates* (Ed. by S. A. Altmann), pp. 281–324. Chicago: University of Chicago Press.
- Zoloth, S. R., Petersen, M. R., Beecher, M. D., Green, S. G., Marler, P., Moody, D. B. & Stebbins, W. C.** 1979. Species-specific perceptual processing of vocal sounds by monkeys. *Science*, **204**, 870–873.
- Zuberbühler, K., Noë, R. & Seyfarth, R. M.** 1997. Diana monkey long-distance calls: messages for conspecifics and predators. *Animal Behaviour*, **53**, 589–604.