

Functional referents and acoustic similarity: field playback experiments with rhesus monkeys

MARC DAVID HAUSER

Departments of Psychology and Anthropology, Program in Neurosciences, Harvard University

*(Received 3 March 1997; initial acceptance 22 May 1997;
final acceptance 2 October 1997; MS. number: A7871)*

Abstract. Rhesus monkeys, *Macaca mulatta*, on the island of Cayo Santiago, Puerto Rico produce one or more of five acoustically distinctive calls when they find food. Three of these calls ('warbles', 'harmonic arches' and 'chirps') are produced by individuals finding high-quality, rare food items, whereas the other two calls ('coos' and 'grunts') are produced upon encountering lower-quality, common food items, and in non-food contexts as well. To determine how rhesus classify such acoustic variation, I conducted habituation experiments using a subset of the five call types. I designed experiments to reveal whether classification is based primarily on acoustic features or on the basis of a call's functional referent; caller identity was held constant within sessions. Habituation to 'warbles' transferred to 'harmonic arches', and vice versa. Thus, although these two calls are acoustically distinctive, they appeared to be perceptually clustered into one category based on referential similarities. In contrast, habituation to 'grunts' was followed by dishabituation to 'warbles' or 'harmonic arches', and habituation to 'warbles' or 'harmonic arches' was followed by dishabituation to 'grunts'. Dishabituation could be due to acoustic or referential differences. Significantly, the magnitude of the dishabituating response was asymmetric and depended upon the call type used in the habituation series. Thus, when subjects were habituated to 'grunts', they responded significantly more to tests of 'warbles' or 'harmonic arches' than when the sequence was reversed. These results suggest that for rhesus monkey food-associated calls, referential differences carry more weight during perceptual classification than do acoustical differences.

© 1998 The Association for the Study of Animal Behaviour

When stimuli are perceived, they are typically classified into functionally meaningful categories to facilitate response selection. Membership within a functional category will depend, in part, on a set of salient features. What provides coherence among exemplars of a category is that they share certain features. An important goal of work on non-human animal communication is to extract such features, thereby providing a window into the mechanisms underlying perceptual classification.

In recent years, observations and experiments have raised the possibility that monkeys and birds classify vocalizations on the basis of their meaning (Cheney & Seyfarth 1990; Marler et al. 1992; Hauser 1996). The main idea here is that like some human words, non-human animal vocalizations

are functionally referential in that they pick out particular objects and events in the environment. Such claims do not deny the importance of affective information in the signal (e.g. Smith 1977; Marler 1985; Owings 1994). Rather, they suggest that when call exemplars are classified into functional categories, the call's putative referent represents a common factor (Dittus 1984; Gouzoules et al. 1984; Marler et al. 1986; Macedonia 1991; Hauser & Marler 1993a; Evans & Marler 1994; Evans et al. 1994).

Evidence of functionally referential signals comes from three types of research: (1) detailed observations of the contexts in which signals are produced, with evidence of a tight correlation between signal structure and the putative referent or associated context (e.g. food: von Frisch 1967; Dittus 1984; Marler et al. 1986; Gould 1990; Elowson et al. 1991; Benz 1993; Hauser & Marler 1993a; alarm: Struhsaker 1967; Seyfarth et al. 1980b; Evans et al. 1994); (2) call production

Correspondence: Marc D. Hauser, Departments of Psychology and Anthropology, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138, U.S.A. (email: hauser@wjh.harvard.edu).

experiments involving manipulations of the putative referent (e.g. food: von Frisch 1967; Gould 1990; Hauser 1992; Benz 1993; Hauser & Marler 1993b; Evans & Marler 1994; alarm: Pereira & Macedonia 1991); (3) perceptual experiments involving playbacks of naturally recorded vocalizations (e.g. alarm: Seyfarth et al. 1980a; Cheney & Seyfarth 1988, 1990, 1992; Macedonia 1991; Evans et al. 1994). Studies of call production provide information on the range of stimuli eliciting calls, and the extent to which call structure varies in such contexts. Studies of call perception indicate how acoustic variation is classified into functionally meaningful categories and thereby provide insights into the signal's informational content. Far less work has been carried out on call perception (but see Moody et al. 1990; Stebbins & Sommers 1992), and this is especially the case for primates living under natural conditions.

To date, two types of perceptual experiments have been used in the study of functionally referential signals. The first attempts to demonstrate whether the acoustic features of a signal are sufficient to elicit a behaviourally appropriate response. This experimental design depends upon well-defined response assays that enable a distinction between call types. Thus, for example, in Seyfarth et al.'s (1980a) experiments with vervet monkey, *Cercopithecus aethiops*, alarm calls, three types of response were defined on the basis of observations during natural predator encounters (Struhsaker 1967; Seyfarth et al. 1980b): searching the ground in response to snake alarms, scanning the sky and then running into a bush in response to eagle alarms, and running up and on to high, thin tree branches in response to leopard alarms. The second approach uses a habituation procedure to contrast two possible classification mechanisms. When calls are perceived, they can be classified on the basis of their acoustic properties alone or in terms of a conceptually mediated property such as the call's referent. Cheney & Seyfarth (1988) were the first to explore this paradigm. They found, for example, that following habituation to one type of inter-group call, vervet monkeys transferred habituation to the second, acoustically different, type of inter-group call, as long as caller identity was the same in the habituation and test trials. Cheney & Seyfarth concluded that although the two inter-group calls are acoustically different, vervets classify them into one functionally meaningful category.

One potential problem for the habituation paradigm arises when the habituation and test stimuli have both acoustic and referential differences. Here, dishabituation to the test stimulus could be due to either factor or both. A possible way out of this dilemma is to determine whether there are asymmetries in the magnitude of dishabituation contingent upon the order of stimuli presented during habituation (Tversky 1977). For example, after hearing a customer in a restaurant say, repeatedly, 'Yum, potatoes!', one would eventually habituate. If the customer then announced 'Yum, caviar!' one would certainly respond, both because 'caviar' refers to something different from 'potatoes' and because it sounds different. Now imagine that the customer started by repeating 'Yum, caviar!' and then switched to 'Yum, potatoes!'. Here, although there are both acoustical and referential differences (the same ones that we have in our first example), one might ignore the change or simply give it a quick look. Although 'caviar' and 'potatoes' are being used in two different test sequences, there is an asymmetry in response due to the perceptual and motivational state that results following habituation to the first repeated utterance. And such differences in state can only be elicited if one is sensitive to a functional difference between the utterances, one driven by the referential differences between potatoes and caviar. This logic was used to evaluate how rhesus monkeys, *Macaca mulatta*, classify a subset of their food-associated calls.

When rhesus monkeys living on the island of Cayo Santiago, Puerto Rico find food, they give one or more of five acoustically distinctive calls, hereafter referred to as food-associated calls (Hauser & Marler 1993a). Three calls ('warble', 'harmonic arch', chirp) are given by individuals that find high-quality, rare food items (e.g. coconut). Two other calls ('coos', 'grunts') are given to lower-quality, common food items (e.g. monkey chow), but are also given in non-food contexts (e.g. group movement, grooming). These food-associated calls can be discriminated by human listeners and presumably by the animals themselves. Moreover, they can be differentiated on a wide number of spectral and temporal parameters (Hauser & Marler 1993a). Observations and experiments also reveal that call rate co-varies with the individual's hunger level, whereas differences in the acoustic morphology of the call co-vary with the type of food discovered. This

suggests that the affective component of these calls is carried by a different acoustic vehicle than the functionally referential component. Although subjects respond to food-associated calls by looking, and sometimes moving in the direction of the caller, we have no evidence that rhesus respond differently to different call types. The experiments described below, therefore, use the amount of time spent looking in the direction of the speaker as an assay to assess the mechanisms guiding call perception.

METHODS

Subjects and Study Area

I studied semi-free-ranging rhesus monkeys living on the island of Cayo Santiago, Puerto Rico (for general description of the island and population, see [Rawlins & Kessler 1987](#)). At the time of our observations, the population consisted of approximately 900 individuals, separated into six social groups. The population has been observed for almost 60 years. Consequently, detailed demographic, reproductive and behavioural data are available. The animals are well habituated to human observers, and have had experience with a variety of experimental situations, including playbacks ([Gouzoules et al. 1984](#); [Hauser & Marler 1993b](#); [Hauser & Andersson 1994](#); [Rendall et al. 1996](#)).

Rhesus monkeys on Cayo Santiago live in multi-male, multi-female social groups, as is characteristic of rhesus living in more natural situations ([Melnick & Pearl 1987](#); [Fa & Lindburg 1996](#)). At approximately 0730 hours each day, Purina monkey chow is placed into the three dispensers on the island. The chow is typically depleted by approximately 1400 hours each day. In addition to eating the provisioned chow, individuals also eat fruits, flowers, insects, leaves, grass and soil available on the island. The non-provisioned food makes up between 35 and 70% of the diet; variation arises due to season and a group's dominance. For example, lower-ranking groups are often supplanted at the dispensers by higher-ranking groups and thus are forced to forage on naturally available food items. There are no predators on the island. Mortality among individuals arises due to old age, injuries from fights and lack of consistent access to the chow dispensers.

All subjects tested were fully adult. For rhesus monkeys on Cayo Santiago, females reach reproductive maturity at approximately 3 years and males at approximately 4 years ([Bercovitch & Berard 1993](#)).

Recordings and Playbacks

I used focal-animal and all-occurrence sampling techniques to obtain recordings of food-associated calls and their contexts ([Hauser & Marler 1993a](#)). I recorded vocalizations using a Sennheiser MKH-816 directional microphone and either an analog (Sony TCD-5M) or digital cassette-recorder (Sony TCD-D7); this recording system is sufficient to capture the full bandwidth of rhesus monkey vocalizations (100 Hz–16 kHz) and has been used in several other studies ([Hauser 1991, 1993](#); [Hauser et al. 1993](#)). All vocalizations were recorded at a distance of 1–3 m from the subject. Once acquired, signals were digitized (25–40 kHz sample rate) onto a portable computer using the SIGNAL sound analysis system ([Beeman 1996](#)). Signals were inspected spectrographically and edited for playbacks. Editing included cropping the signal to eliminate background noise or artificial pops at onset and offset. If noise from the wind or surf did not overlap with the spectral range of the signal, it was digitally filtered. This process provided a set of playback stimuli with relatively high signal-to-noise ratios. I then imported digitally acquired signals into a software interface which included a list of calls, a volume control for sound output, and play-stop commands for initiating or terminating a playback.

I implemented a habituation paradigm to determine how rhesus monkeys classify their food-associated vocalizations. The general technique has been used under laboratory conditions, especially with human infants ([Eimas et al. 1971](#); [Spelke 1985](#)). It has been used less frequently with non-human animals ([Wytenbach et al. 1996](#)), especially under natural conditions ([Petrinovich & Patterson 1979, 1981](#); [Cheney & Seyfarth 1988](#); [Nelson & Marler 1989](#); [Rendall et al. 1996](#)).

The experimental design was as follows. Each playback session included a series of habituation trials followed by a test trial. In order for a session to be included in the final analyses, the subject had to respond on at least the first two trials of the habituation series; we discarded 11 sessions in

which the subject failed to respond on the first or second habituation trial. Habituation trials were continued until the subject failed to respond on two consecutive trials. This criterion places all subjects in a comparable attentional state during the test trial. Specifically, all subjects entered the test trial after demonstrating, on two trials, that they were no longer overtly interested in the call type used during the habituation series. By using a behavioural criterion to determine habituation, this procedure is comparable to the habituation experiments used by many developmental psychologists to assess what pre-linguistic infants know about the physical world (Spelke 1985; Hauser & Carey, *in press*), in addition to numerous experiments in animal learning (Brogden 1951; Hilgard 1951; Petrinovich 1981; Thompson *et al.* 1973).

On completion of their habituation trials, I presented the subjects with a single test stimulus from a different acoustic category, but given by the same vocalizer. Inter-trial interval was set between 5 and 60 s following the subject's response. The minimum was based on intervals between naturally occurring food-associated calls, which were never less than 5 s. Although the interval between successive food-associated calls can exceed 60 s under natural conditions, I restricted it to increase the probability that subjects would stay in the same location for the entire playback series, and remain in a comparable motivational/behavioural state.

If subjects failed to respond to the test stimulus, I played a post-test signal from a functionally different category, a 'shrill bark', given by individuals that were alarmed (Bercovitch *et al.* 1995). The post-test trial was designed to examine whether failure to respond in the test trial was due to overall habituation to the playback site, rather than to the test stimulus. In five sessions, subjects failed to respond to the shrill bark. I therefore assumed that they had habituated to the test situation and eliminated these sessions from the analyses.

I conducted playback experiments using a portable computer and an Anchor AN-256 speaker (frequency range=55 Hz–18 kHz; 55 W output). This system has been used in previous playback experiments (Hauser & Andersson 1994). Playback trials were initiated using the following procedures and conditions. (1) Test subjects were sitting in a relatively open area with relatively few

neighbours (<10 within a 10-m radius). (2) The speaker was placed in dense vegetation, presumed to be out of sight from the subject, and situated approximately 15–25 m away. When the speaker was in place, I waited 5 min before starting the experiment to minimize association with the playback equipment; to remove further association with the playback equipment, I periodically placed speakers in bushes, set up for the experiment, but failed to initiate the trial. One observer was responsible for running the computer, 10–15 m away from the speaker. A second observer videotaped the subject's response. (3) Sessions were aborted if inter-trial intervals exceeded 1 min. This occurred when subjects moved, when a food-associated call (identified by ear) was heard within 30 min of the initiation of the playback, or when a fight broke out, disrupting the test area. Sessions were also aborted if the available number of unique call exemplars within a call type was less than the number of trials needed to habituate a subject; this occurred twice. For example, one 'harmonic arch' series consisted of 10 exemplars. During one session, a subject responded on the tenth habituation trial. Thus, it was not possible to continue the session without re-using an exemplar that had already been played.

Response to playback was defined as a distinctive head turn, followed by a look towards the speaker (head rotates with eyes looking in the direction of the speaker) within 2 s of stimulus onset; although some subjects may have turned their eyes towards the speaker without turning their head, it was not possible to detect this in the field nor from our video footage. Because the speaker was placed at a 90° angle from the subject, judging a response was relatively unambiguous. In cases where the response was ambiguous, however, I flagged the trial on a note pad and then re-evaluated it from the video record (see below).

Videotapes were transcribed using Radius' Video Vision digitizing board, a Macintosh and the Adobe Premiere software. This system provides digitized records of the audio and video tracks, which are subsequently analysed in the following steps. First, all sessions are saved as files and assigned a number that corresponds to a master list indicating the playback stimuli used. When a session is analysed, however, there is no information about call type, and thus all trials are scored blind. Second, a visual representation of

Table I. Experimental stimuli and test conditions

Playback condition	Habituating stimulus	Test stimulus	No. subjects
Within-referent	'Warble'	'Harmonic arch'	6
Within-referent	'Harmonic arch'	'Warble'	6
Between-referents	'Warble'/'harmonic arch'	'Grunt'	7
Between-referents	'Grunt'	'Warble'/'harmonic arch'	7

the digitized video and audio (time-amplitude waveform) records are simultaneously displayed on the computer monitor. Third, the observer moves through the audio track and finds the onset of the signal from the waveform, and then flags the corresponding video frame. Fourth, the amount of time subjects spent looking towards the speaker was scored using frame-by-frame (30 frames/s) analyses. In cases where there was disagreement between the field observer's assessment of a response and the video scorer's assessment, the session was thrown out; this caused us to reject 26 sessions, but allowed us to evaluate trials with perfect inter-observer reliabilities with respect to the presence or absence of a response.

To evaluate patterns of response across call types, and assess whether natural differences in response to each call type emerge on the first playback trial in contrast to subsequent trials, I evaluated the average response per trial within the habituation series up to and including trial 5; the sample size for trials 6–10 was less than five, which is insufficient for statistical analyses.

I used parametric *t*-tests, justified by analyses of normality, to compare response durations and habituation slopes. I used non-parametric sign tests to evaluate the number of individuals dishabituating in the test trials.

Experimental Stimuli and Conditions

Using the general experimental protocol described in the previous section, subjects were tested in one of two broad conditions: within-referent and between-referents. Playbacks focused on three call types: 'warbles', 'harmonic arches' and 'grunts' (Fig. 1, Table I).

For the within-referent condition, eight males and four females were tested. For the between-referents condition, 10 males and four females were tested. All subjects were members of the

same social group and no individual took part in more than one session.

Within a playback session, calls from one individual were broadcast and this individual was a

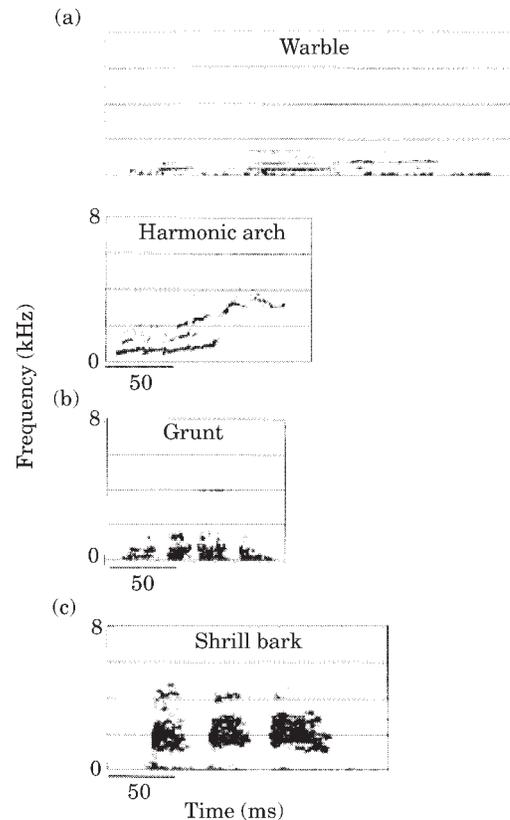


Figure 1. Sound spectrograms of stimuli used in field playback experiments with rhesus monkeys. The 'warble' and 'harmonic arch' are given by individuals that have discovered (a) rare food of high quality. (b) The 'grunt' is given by an individual that has discovered a common food of lower quality. (c) The shrill bark is an alarm call and was used as a post-test signal. The Y-axis represents frequency (kHz) and the X-axis represents time (ms).

member of the test subject's social group. For each call type, there were two unique series of exemplars, each series from a different individual. Within each playback condition, there was an even distribution of each of the two playback series. By restricting the exemplars in a session to one individual, we avoided confounding caller identity and call type. Prior to presentation, a random number generator was used to determine the sequence of exemplars within a habituation series, as well as the exemplar to be used in the test trial. Consequently, each subject received a different order of exemplars in both the habituation and test trials. Because the total number of trials to habituation varied between subjects (see Results), no single call exemplar appeared disproportionately in the penultimate or final habituation trials.

The two referent conditions were set up on the basis of prior observations and experiments, described in the Introduction. In particular, both 'warbles' and 'harmonic arches' are produced by individuals that have found rare, high-quality food items, whereas 'grunts' are produced by individuals that have discovered commonly available food such as monkey chow. These call types can be differentiated on the basis of both spectral and temporal parameters (Hauser & Marler 1993a; Fig. 1). For example, 'warbles' are longer in duration than 'harmonic arches' which are longer than 'grunts'. 'Harmonic arches', have a higher fundamental frequency than 'warbles', which have a higher fundamental frequency than 'grunts'. These contextual and acoustic differences lead to the following predictions.

Acoustic hypothesis

If the three food call types are classified on the basis of acoustic differences alone, then following habituation to one call type, playbacks of either of the other call types should result in a detectable response (i.e. failure to transfer habituation) because they are acoustically distinct.

Referential hypothesis

If the three food call types are classified on the basis of referential properties, then following habituation to 'warbles', subjects should transfer habituation to 'harmonic arches' because they share a common referent; this pattern should also

hold when subjects are habituated to 'harmonic arches' and then tested with 'warbles'. This prediction is based on the observation that 'warbles' and 'harmonic arches' are produced in the same general context, finding rare, highly preferred food. However, when subjects are habituated to either 'warbles' or 'harmonic arches' and then tested with 'grunts', they should show a significant response to the 'grunt', which is produced in a functionally different context; this pattern is also expected when subjects are habituated to 'grunts' and tested with either 'warbles' or 'harmonic arches'.

A problem for the referential hypothesis is that dishabituation in the test trial of the between-referents condition could result from either acoustic or referential differences. That is, 'grunts' are acoustically different from 'warbles' and 'harmonic arches', and may also be referentially different. To disentangle these two possibilities, we can look for asymmetries in the strength of the response to the test stimulus based on the call type used during habituation. Thus, if 'warbles' represent a stronger and more interesting call than 'grunts' (i.e. in terms of the food discovered), then following habituation to 'warbles', we would expect either no response or a weak response to a 'grunt'. In contrast, following habituation to 'grunts', we would expect a stronger response to 'warbles'. If differences in the magnitude of response are observed, this would provide stronger support for the referential hypothesis.

RESULTS

Our experimental design was set up to allow subjects to determine the total number of habituation trials presented prior to the test trial. Comparing across call types, the modes (not including the two no-response trials) were comparable ('warbles'=5; 'harmonic arches'=6; 'grunts'=6), but the ranges were slightly different ('warbles'=5-9; 'harmonic arches'=4-10; 'grunts'=5-7).

There was little variability within trials in the magnitude of the response (Fig. 2). Mean response duration was greater on the first trial with 'harmonic arches' ($\bar{X} \pm SD = 6.3 \pm 1.0$ s) than on the first trial with 'warbles' ($\bar{X} \pm SD = 4.6 \pm 1.8$ s; $t_{11} = 13.1$, $P < 0.0001$). Results from a regression comparing slopes (i.e. habituation curves), how-

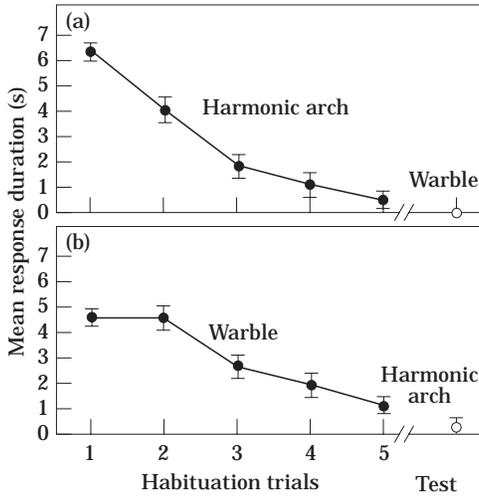


Figure 2. Results from within-referent habituation experiments where subjects were (a) habituated to 'harmonic arches' and then tested with a 'warble' and (b) habituated to 'warbles' and then tested with a 'harmonic arch'. Data points represent the $\bar{X} \pm SD$ time (s) individuals spent looking in the direction of the speaker.

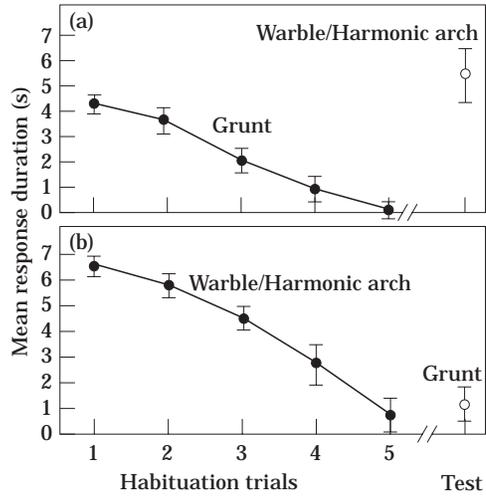


Figure 3. Results from between-referents habituation experiments where subjects were (a) habituated to 'grunts' and then tested with either a 'warble' or 'harmonic arch' and (b) habituated to either 'warbles' or 'harmonic arches' and then tested with a 'grunt'. Data points represent the $\bar{X} \pm SD$ time (s) individuals spent looking in the direction of the speaker.

ever, revealed no significant difference between call types ($t=0.76$, $P>0.05$). By trial 5, there was no statistically significant difference in the magnitude of response to 'harmonic arches' ($\bar{X} \pm SD = 0.5 \pm 0.9$ s) compared with 'warbles' ($\bar{X} \pm SD = 1.2 \pm 0.8$ s; $t_{11} = 1.01$, $P>0.05$). Moreover, and as defined by our experimental design, all subjects entered the test trial after two consecutive trials associated with no detectable response.

Following habituation to either 'harmonic arches' or 'warbles', only one out of 12 subjects responded in the test trial: an adult male, after having habituated to 'warbles', turned and looked towards the speaker for 1.2 s in response to the presentation of a 'harmonic arch'. Thus, a statistically significant number of individuals transferred habituation (sign test: $P<0.01$). All subjects that failed to respond in the test trial responded in the post-test trial.

The mean duration of response on the first habituation trial to 'grunts' ($\bar{X} \pm SD = 4.2 \pm 2.2$ s) was significantly less than the mean duration of response to 'warbles' or 'harmonic arches' ($\bar{X} \pm SD = 6.6 \pm 1.2$ s; $t_{13} = 9.39$, $P<0.0001$; Fig. 3). Thus, the spontaneous response (i.e. in the absence of a preceding stimulus presentation) to 'warbles' and 'harmonic arches' was consistently

greater than it was to 'grunts'. Results from a regression comparing slopes revealed no significant difference in habituation rates ($t=1.46$, $P>0.05$). By the fifth habituation trial, there was no statistically significant difference between responses to 'grunts' ($\bar{X} \pm SD = 0.1 \pm 0.3$ s) compared with 'warbles' or 'harmonic arches' ($\bar{X} \pm SD = 0.8 \pm 0.9$ s; $t_{13} = 1.12$, $P>0.05$).

Following habituation to 'grunts', all seven subjects responded to the test trial of 'warbles' or 'harmonic arches'. A one-sample t -test with a hypothesized mean of 0.0 (i.e. subjects that failed to respond on two consecutive trials scored a zero response) revealed a statistically significant difference in response in the test trial ($t=9.04$, $P<0.0003$). Following habituation to 'warbles' or 'harmonic arches', six out of seven subjects responded to 'grunts'. A one-sample t -test with a hypothesized mean of 0.0 revealed a statistically significant difference in response in the test trial ($t=4.90$, $P<0.003$). Subjects' responses to the 'warble'/'harmonic arch' test were also greater than their responses to the 'grunt' test ($t_{12} = 7.39$, $P=0.0001$). Considering all subjects tested in the between-referents condition, 13 of 14 responded (i.e. failed to transfer habituation) in the test trial (sign test: $P<0.001$).

DISCUSSION

To understand better the factors guiding classification of acoustic signals in rhesus monkeys, we set out to determine whether individuals classify a subset of their food-associated calls on the basis of salient acoustic features or in terms of their referents. Subjects showed a stronger response to 'harmonic arches' than to 'warbles' on the first trial of the habituation series and consistently transferred habituation from the final trial of the habituation series to the test trial. Although subjects failed to dishabituate in the test trial, they subsequently responded to the signal ('shrill bark') presented in the post-test trial, suggesting that they were not merely habituating to the playback situation or spatial location of vocal activity. Although 'warbles' and 'harmonic arches' are acoustically different, subjects appeared to be clustering them into one category. From a functional perspective (Cheney & Seyfarth 1988, 1990), these playbacks simulate an animal that unreliably calls about the presence of an object or event, food in the current situation. Thus, when a rhesus monkey repeatedly signals the presence of food with 'warbles', a shift to 'harmonic arches' fails to signal a functionally significant shift in the putative referent or context-eliciting condition.

In our between-referents condition, subjects consistently showed a stronger response on the first trial to 'warbles' and 'harmonic arches' (continuous, broad band, tonal signals) than to 'grunts' (a pulsatile, broad band, atonal signal). Following habituation to one call type, subjects consistently dishabituated to the other call type. This result failed to reveal whether rhesus classify these calls on the basis of acoustic or referential differences since both factors can contribute to dishabituation. When we compared response magnitudes in the test trials, a striking difference emerged. When subjects were habituated to 'grunts', the magnitude of the response to 'warbles' or 'harmonic arches' was always greater than the response to 'grunts' on the first habituation trial, despite individual differences. In contrast, when subjects were habituated to 'warbles' or 'harmonic arches', the magnitude of the dishabituating response to a 'grunt' was small, compared with that on the fifth trial of the habituation series. Again, although inter-individual variation was observed, the magnitude of the response was small for all individuals, typically of the order

of 1.5 s. This result suggests that in terms of their effects on the orienting response, 'warbles' and 'harmonic arches' are inherently stronger stimuli than 'grunts'. The fact that individuals dishabituated to 'grunts' following habituation to either 'warbles' or 'harmonic arches', but not vice versa, lends additional support for the hypothesis that calls are being classified with respect to their referents.

Returning to the functional perspective sketched above, when a rhesus monkey repeatedly 'grunts', a shift to 'warbles' or 'harmonic arches' represents a significantly more potent message for it appears to indicate a shift to a different class of food, something that is rare and of high quality. When this situation is reversed, the shift in call types is significant, but the message is less potent. Returning to our hypothetical example in the Introduction, if you have been eating caviar for an hour, you are unlikely to be tempted or interested by someone announcing a second course of potatoes. In contrast, if the first course is potatoes, you will surely be tempted and interested by the announcement of caviar. The acoustic differences are the same in either direction, but the referential differences are not.

What alternatives might account for the pattern of results presented? Consider the results from the within-referent playbacks. The difficulty lies in our attempt to account for no response in the test trial. To discount the possibility that subjects habituated to the playback site, we ran a post-test trial at the same location with a call type ('shrill bark') that was both acoustically and referentially different from calls in the habituation and test trial. On every session where subjects failed to respond in the test trial (i.e. transferred habituation), they responded in the post-test trial. Response to the 'shrill bark', however, could arise because it is both referentially different, as well as significantly different in all of its morphological details from food-associated calls. Thus, acoustic contributions to call classification cannot be excluded. Future experiments should use post-test signals that are given in a different socioecological context from the habituation and test signals, but are more similar in acoustic structure (e.g. a 'coo'). Given the consistency of the present results (i.e. lack of dishabituation, low inter-individual variation in response both within and across trials, comparable habituation curves), I suggest that rhesus monkeys classify 'warbles' and

'harmonic arches' into one functional class, calls given to high-quality/rare food items.

Another possible interpretation of the data comes from considering work on categorical perception (reviewed in Harnad 1987; Kuhl 1989; Hauser 1996). Categorical perception occurs when an acoustic continuum is sharply divided into two categories. Critically, exemplars within a category are considered the same, even when physical differences between exemplars within a category are equal to those between exemplars that straddle the boundary between categories. Subjects cannot only discriminate between exemplars in each category, but can also label exemplars from each category. Studies of human infants and several non-human animals reveal that a number of phonemic distinctions (e.g. 'ba' versus 'pa') are perceived categorically, using salient acoustical differences. Consider for the moment that the distinction between 'warbles'/'harmonic arches' and 'grunts' is like the difference between 'ba' and 'pa'. One clear difference between the two rhesus categories is in terms of the pulsatile nature of the signal: 'warbles' and 'harmonic arches' are signals with continuous energy, whereas 'grunts' are pulsatile (i.e. discrete bursts of energy separated by periods of silence). To my knowledge, no study of categorical perception of human speech phonemes has revealed an asymmetry in the magnitude of the dishabituating response, dependent upon the habituating signal (i.e. the dishabituating response is the same following habituation to either a 'ba' or a 'pa'). In contrast, an asymmetry was observed for the rhesus calls, suggesting that something other than the acoustic morphology of the signal is causally relevant to the response pattern. We hypothesize that the crucial factor is the call's referent.

In Cheney & Seyfarth's (1988; see also Rendall et al. 1996) playback design, caller identity was controlled within some sessions and allowed to vary as an experimental condition in other sessions. In the experiments presented here, caller identity was held constant within sessions. For vervet inter-group vocalizations, when caller identity is held constant, habituation to one call type results in subjects transferring habituation to the other call type during the test trial if the two call types were originally given in the same context. In contrast, when caller identity changes between the habituation and test trials, subjects dishabituate in the test trial, even if the two call types were

originally given in the same context. Functionally, although one individual is perceived as unreliably signalling the presence of a neighbouring group, this perception does not transfer to other individuals. Similar experiments could be conducted with the rhesus monkeys' 'warble' and 'harmonic arch' to determine whether subjects dishabituate when caller identity changes from the habituation to the test trial.

The between-referents condition was also associated with considerable consistency across individuals: only one subject out of 14 failed to dishabituate in the test trial, and responses to the test signal were relatively consistent across subjects, with all individuals showing a stronger response to the 'warble' or 'harmonic arch' than to the 'grunt'. Although caller identity was also held constant in the between-referents condition, I expect that changing it would simply increase the overall magnitude of the response in the test trial.

Two questions remain: why do rhesus have two, or possibly three (i.e. the chirp) calls for the same putative referent and what is the precise meaning or referent of these calls? The 'warble' and 'harmonic arch' are both given by an individual that has discovered rare, high-quality food. At present, we have not found differences in the types of food associated with 'warbles' as opposed to 'harmonic arches'. On a coarse-grained level, then, this part of the rhesus monkeys' food-associated call system may be like the vervet monkeys' inter-group call system: two acoustically different calls given in comparable situations. Paralleling the results from the vervets' inter-group calls, however, there may be differences in the affective state of the animal producing 'warbles' as opposed to 'harmonic arches'. Thus, 'harmonic arches' may be given by animals that appear relatively more excited and hungry than individuals giving 'warbles'. Many of the 'harmonic arches' that have been recorded were produced by animals that, during the trapping season, were extremely hungry (the provisioned food is withheld for a few days) and had discovered corn, a food item that is only put out during trapping (i.e. once a year) to lure hungry subjects into the corals. Moreover, our experimental results revealed that subjects responded more intensely on the first trial of the habituation series to 'harmonic arches' than to 'warbles'. To assess the validity of this speculation about affective state, long-term studies will be needed due to the

infrequency with which 'warbles' and 'harmonic arches' are produced. Additionally, it might be possible to conduct more manipulative experiments in captivity, where hunger level and food quality can be controlled better and then explored for their effects on food-calling behaviour.

Gaining a complete understanding of the precise meaning of these food-associated calls is extremely difficult. In contrast to the studies of vervet alarm calls, the behavioural response to food calls is typically an orienting response in the direction of the caller; occasionally, subjects move a few steps in the direction of the caller. In the case of vervet monkey alarm calls, although there are several possible meanings, the subject's response provides a helpful guide. Thus, when vervet monkeys hear an alarm call given in response to an eagle, they scan up, whereas when they hear an alarm call to a snake, they stand bipedally and scan down. Because such scanning is oriented towards different spatial targets, it must be for something, and something different in each case. The putative referent could be as general as 'dangerous thing in the air' as opposed to 'dangerous thing on the ground' or it could be more specific. The specificity of the vervets' response to alarm calls provides a tool for refining the specificity of the referent in a way that is much more difficult with a general orienting response.

It is also difficult to determine whether the call's referent is a label for an object or event (e.g. 'a leopard', 'a rare food item'), or whether it is a command for action (e.g. 'run up into a tree to escape danger', 'come over here to eat'); see Marler (1961) and Cheney & Seyfarth (1990) for a discussion of this point. As suggested elsewhere (Hauser 1996), distinguishing between these alternative accounts will require new experimental procedures. For example, to test whether rhesus food-associated calls are labels or invitations to approach, one could contrast the calling behaviour of individuals in a two-member group that are alone and eating as opposed to eating with the other group member. Calls by the lone individual could be interpreted as labels or invitations. If calls are given when both group members are present and eating, however, then an invitation to feed can be ruled out. There is, in the Quinean (Quine 1973) sense, a potential sea of meanings (e.g. 'stay and eat with me'), but with such tests one can begin to rule out some alternatives.

ACKNOWLEDGMENTS

For comments on the manuscript, I thank D. Cheney, W. T. Fitch, P. Marler, L. Petrinovich, R. Seyfarth, P. Waser and two anonymous referees. For help with data collection and analysis, I thank L. Steffey, P. MacNeilage and M. Ware. For access to facilities on Cayo Santiago, as well as the long-term data base, I thank the staff of the CPRC, funding from the NIH (USPHS Grant CM-5-P40-RR03640) and especially, Drs J. Berard, F. Bercovitch and M. Kessler. Funding for this work was provided by grants from a National Science Foundation Young Investigator Award, the Leakey Foundation, the Wenner-Gren Foundation and by Harvard University.

REFERENCES

- Beeman, K. 1996. *SIGNAL User's Guide*. Belmont, Massachusetts: Engineering Design.
- Benz, J. J. 1993. Food-elicited vocalizations in golden lion tamarins: design features for representational communication. *Anim. Behav.*, **45**, 443-455.
- Bercovitch, F. & Berard, J. 1993. Life history costs and consequences of rapid reproductive maturation in female rhesus macaques. *Behav. Ecol. Sociobiol.*, **32**, 103-110.
- Bercovitch, F., Hauser, M. D. & Jones, J. 1995. The endocrine stress response and alarm vocalizations in rhesus macaques. *Anim. Behav.*, **49**, 1703-1706.
- Brogden, W. J. 1951. Animal studies of learning. In: *Handbook of Experimental Psychology* (Ed. by S. S. Stevens), pp. 568-612. New York: Wiley Liss.
- Cheney, D. L. & Seyfarth, R. M. 1988. Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Anim. Behav.*, **36**, 477-486.
- Cheney, D. L. & Seyfarth, R. M. 1990. *How Monkeys See the World: Inside the Mind of Another Species*. Chicago: The University of Chicago Press.
- Cheney, D. L. & Seyfarth, R. M. 1992. Meaning, reference, and intentionality in the natural vocalizations of monkeys. In: *Topics in Primatology. Vol. 1: Human Origins* (Ed. by T. Nishida, W. C. McGrew, P. Marler, M. Pickford & F. de Waal), pp. 315-330. Tokyo: Tokyo University Press.
- Dittus, W. P. G. 1984. Toque macaque food calls: semantic communication concerning food distribution in the environment. *Anim. Behav.*, **32**, 470-477.
- Eimas, P. D., Siqueland, P., Jusczyk, P. & Vigorito, J. 1971. Speech perception in infants. *Science*, **171**, 303-306.
- Elowson, A. M., Tannenbaum, P. L. & Snowdon, C. T. 1991. Food-associated calls correlate with food preferences in cotton-top tamarins. *Anim. Behav.*, **42**, 931-937.
- Evans, C. S. & Marler, P. 1994. Food-calling and audience effects in male chickens, *Gallus gallus*: their

- relationships to food availability, courtship and social facilitation. *Anim. Behav.*, **47**, 1159–1170.
- Evans, C. S., Evans, L. & Marler, P. 1994. On the meaning of alarm calls: functional reference in an avian vocal system. *Anim. Behav.*, **45**, 23–38.
- Fa, J. & Lindburg, D. 1996. *Evolutionary Ecology and Behaviour of Macaques*. Cambridge: Cambridge University Press.
- von Frisch, K. 1967. *The Dance Language and Orientation of Bees*. Cambridge, Massachusetts: Harvard University Press.
- Gould, J. L. 1990. Honey bee cognition. *Cognition*, **37**, 83–103.
- Gouzoules, S., Gouzoules, H. & Marler, P. 1984. Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Anim. Behav.*, **32**, 182–193.
- Harnad, S. 1987. *Categorical Perception: The Groundwork of Cognition*. Cambridge: Cambridge University Press.
- Hauser, M. D. 1991. Sources of acoustic variation in rhesus macaque vocalizations. *Ethology*, **89**, 29–46.
- Hauser, M. D. 1992. Costs of deception: cheaters are punished in rhesus monkeys. *Proc. natn Acad. Sci. U.S.A.*, **89**, 12 137–12 139.
- Hauser, M. D. 1993. Rhesus monkey (*Macaca mulatta*) copulation calls: honest signals for female choice? *Proc. R. Soc. Lond. Ser. B*, **254**, 93–96.
- Hauser, M. D. 1996. *The Evolution of Communication*. Cambridge, Massachusetts: MIT Press.
- Hauser, M. D. & Andersson, K. 1994. Left hemisphere dominance for processing vocalizations in adult, but not infant rhesus monkeys: field experiments. *Proc. natn Acad. Sci. U.S.A.*, **91**, 3946–3948.
- Hauser, M. D. & Carey, S. In press. Building a cognitive creature from a set of primitives: evolutionary and developmental insights. In: *The Evolution of Mind* (Ed. by D. Cummins & C. Allen). Oxford: Oxford University Press.
- Hauser, M. D. & Marler, P. 1993a. Food-associated calls in rhesus macaques (*Macaca mulatta*). I. Socio-ecological factors influencing call production. *Behav. Ecol.*, **4**, 194–205.
- Hauser, M. D. & Marler, P. 1993b. Food-associated calls in rhesus macaques (*Macaca mulatta*). II. Costs and benefits of call production and suppression. *Behav. Ecol.*, **4**, 206–212.
- Hauser, M. D., Evans, C. S. & Marler, P. 1993. The role of articulation in the production of rhesus monkey (*Macaca mulatta*) vocalizations. *Anim. Behav.*, **45**, 423–433.
- Hilgard, E. R. 1951. Methods and procedures in the study of learning. In: *Handbook of Experimental Psychology* (Ed. by S. S. Stevens), pp. 517–567. New York: Wiley-Liss.
- Kuhl, P. 1989. On babies, birds, modules, and mechanisms: a comparative approach to the acquisition of vocal communication. In: *The Comparative Psychology of Audition* (Ed. by R. J. Dooling & S. H. Hulse), pp. 379–422. Hillsdale, New Jersey: Lawrence Erlbaum.
- Macedonia, J. 1991. What is communicated in the antipredator calls of lemurs: evidence from playback experiments with ring-tailed and ruffed lemurs. *Ethology*, **86**, 177–190.
- Marler, P. 1961. The logical analysis of animal communication. *J. theor. Biol.*, **1**, 295–317.
- Marler, P. 1985. Representational vocal signals of primates. In: *Experimental Behavioural Ecology and Sociobiology* (Ed. by B. Hölldobler & M. Lindauer), pp. 211–222. Stuttgart: Gustav Fischer Verlag.
- Marler, P., Dufty, A. & Pickert, R. 1986. Vocal communication in the domestic chicken. I. Does a sender communicate information about the quality of a food referent to a receiver? *Anim. Behav.*, **34**, 188–193.
- Marler, P., Evans, C. S. & Hauser, M. D. 1992. Animal signals? Reference, motivation or both? In: *Nonverbal Vocal Communication: Comparative and Developmental Approaches* (Ed. by H. Papoušek, U. Jürgens & M. Papoušek), pp. 66–86. Cambridge: Cambridge University Press.
- Melnick, D. J. & Pearl, M. 1987. Cercopithecines in multimale groups: genetic diversity and population structure. In: *Primate Societies* (Ed. by B. B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham & T. T. Struhsaker), pp. 121–148. Chicago: The University of Chicago Press.
- Moody, D. B., Stebbins, W. C. & May, B. J. 1990. Auditory perception of communication signals by Japanese monkeys. In: *Comparative Perception: Complex Perception* (Ed. by W. C. Stebbins & M. A. Berkley), pp. 311–344. New York: John Wiley.
- Nelson, D. A. & Marler, P. 1989. Categorical perception of a natural stimulus continuum: birdsong. *Science*, **244**, 976–978.
- Pereira, M. E. & Macedonia, J. M. 1991. Response urgency does not determine anti-predator call selection by ring-tailed lemurs. *Anim. Behav.*, **41**, 543–544.
- Petrinovich, L. 1981. A method for the study of development. In: *Behavioural Development* (Ed. by K. Immelman, G. Barlow, L. Petrinovich & M. Maim), pp. 90–130. London: Cambridge University Press.
- Petrinovich, L. & Patterson, T. L. 1979. Field studies of habituation. 1. Effect of reproductive condition, number of trials, and different delay intervals on responses of the white-crowned sparrow. *J. comp. physiol. Psychol.*, **93**, 337–350.
- Petrinovich, L. & Patterson, T. L. 1981. The responses of white-crowned sparrows to songs of different dialects and subspecies. *Z. Tierpsychol.*, **57**, 1–14.
- Owings, D. H. 1994. How monkeys feel about the world: a review of 'How monkeys see the world'. *Lang. Comm.*, **14**, 15–20.
- Quine, W. V. 1973. On the reasons for the indeterminacy of translation. *J. Phil.*, **12**, 178–183.
- Rawlins, R. & Kessler, M. 1987. *The Cayo Santiago Macaques*. New York: SUNY University Press.
- Rendall, D., Rodman, P. S. & Edmond, R. E. 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Anim. Behav.*, **51**, 1007–1015.
- 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. *Anim. Behav.*, **28**, 1070–1094.
- Smith, W. J. 1977. *The Behaviour of Communicating*. Cambridge, Massachusetts: Harvard University Press.
- Spelke, E. S. 1985. Preferential looking methods as tools for the study of cognition in infancy. In: *Measurement of Audition and Vision in the First Year of Postnatal Life* (Ed. by G. Gottlieb & N. Krasnegor), pp. 85–168. Norwood, New Jersey: Ablex.
- Stebbins, W. C. & Sommers, M. S. 1992. Evolution, perception and the comparative method. In: *The Evolutionary Biology of Hearing* (Ed. by D. B. Webster, R. F. Fay & A. N. Popper), pp. 211–228. New York: Springer-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: The University of Chicago Press.
- Thompson, R. F., Groves, P. M., Teyler, T. J. & Roemer, R. H. 1973. A dual-process theory of habituation: theory and behaviour. In: *Habituation. Vol. 1* (Ed. by H. V. S. Peeke & M. J. Herz), pp. 128–143. New York: Academic Press.
- Tversky, A. 1977. Features of similarity. *Psychol. Rev.*, **84**, 327–352.
- Wytenbach, R. A., May, M. L. & Hoy, R. R. 1996. Categorical perception of sound frequencies by crickets. *Science*, **273**, 1542–1544.