

A Primate Dictionary? Decoding the Function and Meaning of Another Species' Vocalizations

MARC D. HAUSER

Harvard University

Decoding the function and meaning of a foreign culture's sounds and gestures is a notoriously difficult problem. It is even more challenging when we think about the sounds and gestures of nonhuman animals. This essay provides a review of what is currently known about the informational content and function of primate vocalizations, emphasizing the problems underlying the construction of a primate "dictionary." In contrast to the Oxford English Dictionary, this dictionary provides entries to emotional expressions as well as potentially referential expressions. It therefore represents a guide to what animals do with their vocalizations, as well as how they are represented by signalers and perceivers. I begin by a discussion of the unit problem, of how an acoustic space is carved up into functionally significant components leading to a species-specific repertoire or lexicon of sorts. This section shows how little we know about the units of organization within animal vocal repertoires, and how such lack of information constrains our ability to tackle the problem of syntactic structure. In Section III, I review research on the production and perception of vocal signals that appear to be functionally referential. This work shows that several nonhuman primates produce vocalizations that share some of the key properties of reference, but certainly not all; the components that are missing raises questions about their role as precursors to human words. In Section IV, I explore the social uses of vocalizations, assessing whether the signal contains sufficient information for listeners to judge a caller's credibility; ultimately, caller credibility determines how receivers select an appropriate response. Results show that individuals can use calls to assess whether someone is reliable or unreliable, and that such attributes are associated with individuals and particular contexts. I conclude by synthesizing the issues presented and then raise some directions for future conceptual and methodological progress.

Direct all correspondence to: Mark D. Hauser, Department of Psychology and Program in Neurosciences, Harvard University, 33 Kirkland Street, Cambridge, MA 02138.

I. INTRODUCTION

Translation is a notoriously difficult problem. In literature, we often depend on a translator to capture the beauty of a passage in a foreign language. Some authors, such as the French novelist Celine, felt that style, not story telling, was the essence of literature. Consequently, translations were of little use for they failed to capture the elegance of linguistic style. When Mark Twain's short story *The Celebrated Jumping Frog of Calaveras County* was translated into French, he was so appalled that he back translated it. The upshot of this linguistic ping-pong match: virtually no relationship between the original and Twain's translation of the French.

In philosophy, problems of translation arise in the domain of psychosemantics. A classic example is Quine's (1973) imaginary linguist who travels to a foreign land and attempts to decode the language, starting with "gavagai," a sound uttered by a native as a white rabbit runs by. One's first guess is, of course, that gavagai means white rabbit. But, a sea of alternative meanings flood the scene: undetached rabbit parts, Bugs Bunny, a competitor for the tortoise, something to throw into a stew, a white fur ball that is smaller than Mt. Everest and bigger than a snowflake, and on and on. As Quine pointed out, there is an important sense in which the meaning of an utterance is ambiguous, indeterminate, referentially opaque. Hypothesis testing, by means of collecting heaps of empirical data, will do little to help the cause. In this sense, theories of meaning that depend upon a behavioral analysis are doomed to failure.

Other philosophers such as Fodor (1994) recognize the difficulty of translation, but dismiss Quine's pessimistic view of meaning, arguing that it tends to confuse the informational content of a word with its conceptual extension. Thus, for example, whatever it is that "rabbit" refers to can't be the same thing as what "undetached rabbit parts" refer to, even though they share common ground. A rabbit running by is certainly not the same thing as a bunch of rabbit parts running by. In the natural world, the latter is clearly a violation, whereas the former is not. This distinction points the way to differences in conceptual representation. More importantly for the present discussion, Quine's position misses the point that people's heads are generally not flooded with a sea of alternative meanings; a word's referent is often determinate and meaning relatively transparent. Support for such constraints comes from studies of word learning, showing that the process of mapping from word to referent is constrained in the child by such factors as object knowledge, syntactic cues, and the capacity to attribute intentions to speakers (see, for example, Markman, 1990; Baldwin, 1995; Bloom & Markson, 1998; Carpenter, Nageli, Tomasello, 1998).

Although we may not be happy with a translation, and may find a person's use of a term to be somewhat opaque, we are in a far more difficult situation when it comes to decoding an animal's utterance. There are at least three reasons for this. First, translation, *sensu strictu*, involves moving between two natural languages, where the meaning of words in each language is known. The trick, therefore, involves finding words in one language (e.g., French) that can be matched to words in another language (e.g., English) on the basis of shared referents. In a strict sense, therefore, we cannot translate animal vocalizations

because we lack the corresponding dictionaries. What we can do, however, is determine the function of animal vocalizations by describing how they influence the behavior of an audience, what motivated their production, and the extent to which the information conveyed is conceptually abstract. Second, there are significant problems associated with assessing the appropriate unit of analysis in animal vocal repertoires. In the same way that a child faces the problem of word segmentation during ontogeny (Aslin, Saffran & Newport 1998; Bloom & Markson, 1998; Cutler & Butterfield, 1992; Pallier, Christophe & Mehler, 1997), the ethologist faces a comparable problem in attempting to extract functional units from the acoustic stream. Which sounds are communicative as opposed to accidental byproducts of movement or other physiological changes (e.g., coughing)? Which sounds communicate information about the caller's emotional state, and which about objects and events in the environment? When a sequence of sounds is produced, does the order in which each sound type appears in the sequence affect the overall meaning or message conveyed? Third, the primary assay for uncovering the function of a vocalization, and the information conveyed, is how individuals respond to the vocalization. Sometimes, the behavioral response is dramatic, including displays of aggression, alarm, and sexual machismo. Often, however, the response is subtle, involving at best a nod or focused stare. Such subtlety amplifies the difficulty of the decoding problem. Together, these three problems constrain what the ethologist can deduce about the function of animal vocalizations, as well as the information conveyed.

In this essay, I review what is currently known about the informational content and function of primate vocalizations, emphasizing the problems underlying the construction of a primate dictionary. What I mean here by primate dictionary, however, is not in the usual Oxford English Dictionary (OED) or Webster sense. Rather, the primate dictionary provides entries to emotional expressions as well as potentially referential expressions. It provides a guide to what animals do with their vocalizations, and how they are represented in the minds of signalers and perceivers. Part one focuses on the unit problem, of how particular species carve their acoustic space into functionally significant components, culminating in a species-specific repertoire or lexicon of sorts. Part two reviews work on the production and perception of vocal signals that appear to be functionally referential. Section IV looks at the social uses of vocalizations, exploring whether there is sufficient information in the signal to judge the caller's credibility, and thereby determine the most appropriate response. Section V synthesizes the issues presented and raises some directions for future conceptual and methodological progress.

Two caveats before we begin. First, there is now a rich literature on the production and comprehension by animals of either natural or artificial human languages (Gardner, Gardner & Van Cantfort, 1989; Herman, Patk & Palmer, 1993; Savage-Rumbaugh, 1986; Savage-Rumbaugh, Murphy, Sevcik, Brakke, Williams & Rumbaugh, 1993; Schusterman, Gisiner, Grimm & Hanggi, 1993). Although this work has provided considerable insights into the capacity for acquiring referential signals, I focus here on what we can learn from the natural vocalizations of animals. Second, there are important parallels between the work discussed in this essay and studies of child language acquisition (see, for example, Bloom & Markson, 1998; Hauser, 1996; Seyfarth & Cheney, 1997). Although I make

reference to this work, I do not discuss it in detail. My primary aim is to showcase what we know about primate vocalizations so that researchers working on humans can better appreciate the parallels, both conceptual and methodological.

II. SEARCHING FOR FUNCTIONAL UNITS

“. . . variation in the patterning of natural bird songs is often highly constrained and . . . many learned songs display species-specific universals in many aspects of their structure. Those aspects that do vary can be permuted in many different combinations, providing ample acoustic vehicles for local dialects and individual differences. Other aspects of song structure . . . are shared by all species members, each individual employing some fraction of the universal lexicon.” (Marler, 1997, p. 11).

Unlike birdsong and human speech, we know relatively little about the psychoacoustic units underlying primate vocal communication. To some, this will come as an odd claim. After all, isn't it clear that primates have an array of calls that they deploy during social interactions and predator encounters? Primates certainly use calls to manipulate the behavior of potential listeners and in this sense, the call functions as a significant unit in their vocal repertoire. But what is a call? At one level, it is a segment of acoustic energy that, on its own, plays a functional role in mediating social interactions with conspecifics (e.g., signals of submission or dominance) or heterospecifics (e.g., predator alarm calls). At a different level, calls are part of a system of communication, where any given call is functionally defined in terms of its relationship to other calls. But are there other acoustic levels? In particular, do primates (or any other animal) have units below the level of the call, units that on their own play no functional role, but can be combined with others to create a functionally significant signal? Paralleling human language, do animal vocal repertoires consist of a phonological level? And, at a second level, do primates (or any other animal) string together calls, such that the function or meaning of the string is significantly different from the calls that comprise it? Paralleling human language, do animal vocal repertoires consist of a syntactic level? Taking both levels together, do animal vocal repertoires exhibit the kind of hierarchical structure that is definitional of all human languages?

Consider the first level. Human words are composed of syllables—sequences of consonants and vowels. Syllables, in and of themselves, tend to be meaningless. But the power of our language comes from our capacity to take meaningless syllables and combine them into an unbounded number of meaningful words, and then take these words and combine them into an unbounded number of meaningful expressions (Chomsky, 1986; Studdert-Kennedy, 1998). This conceptual power is derived from what William Abler (1989; 1997) has called “the particulate principle,” the idea that the unlimited diversity of structure within a given system emerges from the capacity to recombine a finite set of discrete particles or elements (atoms in chemistry, integers in our number system, and genes in genetics); the basic elements or particles in these systems have no *meaning* on their own, which is why they provide the basis for a powerful combinatorial system. In contrast, blending systems fail to provide such diversity because the elements

lose their identity or integrity in the same way that white paint and red paint lose their identity when combined to make pink paint. Studdert-Kennedy (1998) has argued that the only way to develop a communication system with the unbounded scope or expressive power of language is by tapping into a particulate system. One challenge for those studying animal systems is, then, to determine whether there are basic acoustic elements that lack meaning on their own, but that when combined with other elements, emerge with meaning.

Bird Song

The songs of many Passerine birds consist of a finite set of notes (discrete bursts of energy) that are combined into syllables (homogeneous string of note). Syllables are then combined into phrases (a segment of different syllables), that are then combined to create novel song types (Marler, 1977, 1997; reviewed in Catchpole & Slater, 1995). Notes, syllables, and phrases lack functional significance; on their own, they play no role in the life of a songbird. The specific structure of the song, in contrast, provides significant information about the singer's identity, local dialect, species, and in some cases, the singer's genetic quality. Three kinds of studies provide support for the validity of these psychoacoustic units in songbirds: development of the repertoire, perceptual classification, and experimental manipulations of production (for reviews of this literature, see Catchpole & Slater, 1995; Kroodsma & Miller, 1996; Snowdon & Hausberger, 1997).

For the young Passerine songbird to acquire a species-typical song, auditory experience is necessary. As Marler (1997) has argued, it now appears that most songbirds come equipped with innately encoded song types which are then activated by auditory experience, stored in memory, and then winnowed down to a set of songs (sometimes just one) as a function of social interactions with community members (Nelson, 1992). For some species, the encoding includes a universal set of notes as well as a set of rules for how to order them, and how often to repeat them. In the lab, such developmental processes are readily documented by means of tape and social tutoring. Under this kind of experimental regime, one can check the relationship between tutored material and song output, both early on in the subsong phase (similar to babbling in human children) as well as later on when song has crystallized into the species-typical form. Results are clear: Birds acquire a crystallized song that represents both the constraints of the encoded template and the material presented. Birds preferentially take the syllables from their own species-typical repertoire and combine these into a species-typical song. On the perception side, several studies have shown that full-fledged aggression by males, or copulatory solicitations by females, are only elicited by full, species-typical song; song from a bird reared in isolation, or syllables played alone, either produce no response or a weakened one (Searcy, 1992).

When a songbird sings, one view of its production system is that once started, the bird engages a tape loop that plays out until finished. If this is correct, it suggests that song, rather than some other acoustic psychoacoustic unit, is the functionally important, neurally determined unit. To test this idea, Cynx (1990) conducted an experiment using adult zebra finch with fully crystallized songs; developmental studies of this species had already

suggested that the syllable was a significant unit. While a zebra finch was singing, its attention was captured by flashing a strobe light at different points during song production. If the tape loop hypothesis is correct, then the strobe should have no effect; the zebra finch should continue singing until the song is finished. In contrast, if there are other levels of organization within the song, then the strobe light should cause the zebra finch to stop before the song naturally ends. Consistently, subjects stopped at syllabic breaks in the song. For this species, therefore, the syllable is a functional production unit. Studies at the neural level are now beginning to support this hypothesis as well (Doupe & Konishi, 1991; Margoliash & Fortune, 1992).

Together, these results indicate that within the bird's song system there are significant psychoacoustic units other than the song, ones that can be recombined in different ways to create functionally different song variations. As studies have shown, however, recombination is constrained by what appear to be species-specific rules for how elements can be ordered. Furthermore, and in contrast to the recombination of words into sentences by humans, the output of songbird recombination does not change its meaning. Specifically, by reordering elements within a song, the singer imposes changes to the informational content of the signal, changes that indicate identity (species, individual, sex, rearing conditions) and perhaps, reproductive status. But the song's function remains the same, based on the fact that it continues to be used in the same context (i.e., announcing territory ownership and mate availability), guided by the same motivational conditions. Beyond the song, it appears that song bouts (strings of songs separated by periods of silence) may provide information about the singer's motivational state (e.g., willingness to defend a territory) and consequently, may also have an indirect effect on the singer's genetic fitness (Lambrechts, 1992).

Nonhuman Primates

Among studies of primates, most consider the call as the fundamental unit based on observations of call-response associations (e.g., call type A causes Fleeing, B causes Alarm), acoustic analyses (e.g., a discrete, bounded burst of energy that is generally longer than a single glottal pulse) and playback experiments that involve transmission of a single call, rather than a string of calls (Cheney & Seyfarth, 1990a; Clark, 1993; Cleveland & Snowdon, 1981; Fischer 1998; Fitch, 1997; Green, 1975; Hammerschmidt & Fischer, 1997; Harcourt, Stewart & Hauser, 1993; Hauser & Marler, 1993a; Macedonia, 1991; Owren, Seyfarth & Cheney, 1997; Rendall, Rodman & Edmond, 1996; Robinson, 1979; Zuberbuhler, Noe & Seyfarth, 1997; Waser, 1977). Although these analyses show that the call plays a functional role in primate communication systems, we must also explore the possibility of other functional units, following the lead of work on bird song and human speech.

Consider the vervet monkey's vocal repertoire, one of the best studied to date. Struhsaker's (1967) early observations indicated that vervets produce approximately 25 or so discrete call types. These call types were identified by integrating detailed field notes about the context of production (e.g., dominance, intergroup threat, predator alarm,

affiliation, infant distress), the response of potential listeners, and acoustic analyses of call morphology. Thus, for example, Struhsaker observed that vervets produce several acoustically distinctive alarm calls upon encountering different predators (i.e., birds of prey, snakes, large cats, humans, baboons), each with their own unique hunting style. Seyfarth, Cheney & Marler (1980; reviewed in Cheney & Seyfarth, 1990a) followed up on this work, using playbacks to determine whether a single alarm call was sufficient to elicit a behaviorally appropriate response. As most readers will know, the playbacks were sufficient (see section III). What have we learned? Clearly, one alarm call is sufficient to elicit an escape response. Thus, the alarm call plays a functional role in the repertoire, perhaps in the same way that the words "Danger!" or "Snake!" play a function role in human communication. But we are still left with the theoretical possibility that there is functional structure at a lower level than the call, in addition to the possibility that there is functional structure at a higher level such as bout or string of alarm calls of the same type.

Three approaches to the unit problem have been tried with primates, but have met with mixed success. First, results from a cross-fostering study indicated that young learned the acoustic morphology of a heterospecific's vocalization, but learned the entire call rather than components of the call (Masataka & Fujita, 1989); a second study failed to replicate these findings, suggesting instead that acoustic experience plays little to no role in the ontogeny of species-typical vocal morphology (Owren, Dieter, Seyfarth & Cheney, 1992, 1993), a conclusion that has been reached by several other authors working on different species (reviewed in Hauser, 1996; Seyfarth & Cheney, 1997; Snowdon & Hausberger, 1997). Second, studies of interpopulation variation in call structure have hinted at the possibility of dialects, but here too, the evidence is rather mixed. Where populations differ, units within a call type are never recombined (Maeda & Masataka, 1987; Masataka, 1988; Mitani, Hasegawa, Gros-Louis, Marler & Byrne, 1992; Fischer, 1998). Instead, quite subtle changes are imposed on the overall structure of the call. For example, in chimpanzees, analyses of the long-distance 'pant-hoot' vocalization reveal increases in the number of build-up elements or the duration of the climax. In all populations, however, the build-up elements come first, and the climax last (for an example of the pant-hoot, see Figure 1). Finally, several authors have carried out detailed analyses of a species' vocal repertoire and have identified cases where either two discrete call types are combined into a new call type, or components from one call type are recombined with components of another to create a third call type (Cleveland & Snowdon, 1981; Robinson, 1979, 1984). These studies clearly show that at the production level, parts of the repertoire can be combined. But unlike studies of chickadees, where there appear to be strict constraints on the order in which notes are combined (Ficken & Popp, 1992; Hailman & Ficken, 1987; Hailman, Ficken & Ficken, 1987), the rules underlying acoustic organization and reorganization are less clear for primate vocal repertoires. Further, in none of these studies, including those on the chickadee, has the referent of a call or call sequence been well established. In some cases, such as the gibbons, it is clear that rearrangement of syllables in the song is related to identity in much the same way as syllable rearrangement allows individuals to be identified in many songbirds (Mitani & Marler, 1989). In other species,

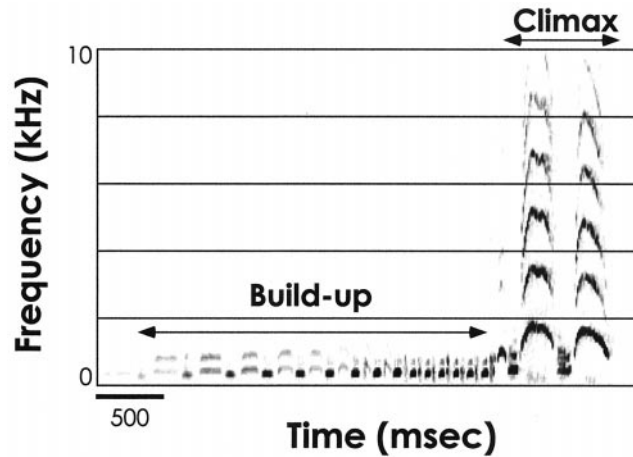


Figure 1. Sound spectrogram of a chimpanzee pant-hoot vocalization, highlighting the build-up and climax components of the call.

such as the capuchin and titi monkey, rearrangement of call types into different sequences maps onto variation in motivational state (e.g., more aggressive or fearful), rather than to the kinds of objects or events referred to.

Having raised some general issues, let me now turn to more specific components of the unit problem in studies of primate vocal communication. I first explore the possibility of units at a level below the call (i.e., syllables) and then examine some evidence for salient information above the level of the call (i.e., a bout)

Foundational Units

A classic approach to understanding the significant features of a signal, is to examine the signal like an anatomist examining a fossilized bone. Break the signal down into features, and then quantify these features by measuring them. Having obtained such measurements, the data are then entered into a statistical procedure such as factor analysis or multidimensional scaling in an attempt to determine which features best distinguish among signal types. With this statistical information in hand, it is then possible to conduct play back experiments to explore how subjects classify signals, and which features of the signal are most salient with respect to classification. For example, studies of the Japanese macaque revealed that for the affiliative “coo” vocalization, changes in the contour or shape of the fundamental frequency (perceptually, the pitch contour) were associated with changes in the social context; thus, there were different coo types associated with group movement, the discovery of food, estrous females, and so forth (Green, 1975). Perceptual experiments revealed that individuals were sensitive to the temporal position of the peak frequency in the fundamental frequency contour, and used this feature for classification; other species did not use the fundamental frequency contour in tests involving perceptual classification of the Japanese macaque coo (Petersen, Beecher, Zoloth, Moody & Stebbins, 1978; Zoloth, Peterson, Beecher, Green, Marler, Moody & Stebbins, 1979). Furthermore, at least

one set of studies showed that Japanese macaques classified coo exemplars categorically, and during such perceptual tasks, showed a left hemisphere bias (Heffner & Heffner, 1984; May, Moody & Stebbins, 1989; Stebbins & Sommers, 1992; for a failure to replicate the categorical perception results, see Hopp, Sinnott, Owren & Petersen, 1992). These studies reveal that particular features within the call determine its morphology, that morphology is associated with a particular function, and that manipulation of morphology outside the normal range of variation transforms the signal from conspecific to something else.

What is unclear about the results just described is the functional work that pitch contours play in the Japanese macaque system. Considering human speech, is the pitch contour a vehicle for emotional expression, comparable to the prosodic cues that we impose on our utterances? Or, is the contour like voice onset time for consonant-vowel pairs, something that allows us to distinguish between “ba” and “pa”? Or, is the contour like a phoneme, something that can be recombined with other sound segments as when we create the words ‘super’ and ‘pursue’ by rearranging the same two syllables. Although the lack of empirical work handicaps our ability to address this problem, let me describe a set of experiments that may open a productive path.

Experiments (Hauser & Andersson, 1994) on the island of Cayo Santiago, Puerto Rico indicate that when adult rhesus monkeys listen to a vocalization played back from a speaker placed 180 degrees behind them, the subject’s right ear leads the orienting response if the call is from the species-typical repertoire. In contrast, if the alarm call of a familiar seabird is played, subjects turn with their left ear leading. This orienting bias has been interpreted as evidence that the left hemisphere is dominant for processing conspecific vocalizations, whereas the right hemisphere is dominant for other significant sounds. In contrast with adult, infant macaques under the age of one year show no orienting bias to either conspecific or heterospecific calls.

Rhesus discriminate conspecific from heterospecific signals. But what acoustic features define “conspecific calls”? Playback experiments were conducted with rhesus in which the temporal structure of three call types was manipulated: a ‘grunt’ given during affiliative interactions, a ‘shrill bark’ given while alarmed, and a ‘copulation scream’ given by males during mating (Hauser, 1998c). Each of these calls consists of a series of pulses (i.e., discrete bursts of energy) separated by brief periods of silence. The manipulation consisted of shrinking the interpulse interval (IPI) to the minimum found in the population or eliminating it completely, or stretching it to the maximum found in the population or twice the maximum; population statistics were determined for each call type based on acoustic analyses of several hundred exemplars. Figure 2 illustrates this signal editing process for the grunt. As in our earlier experiments, we then positioned the speaker 180 degrees behind the target subject and played back a single call. Results showed that for all three call types, subjects turned to the right in response to exemplars with either normal IPIs or the population minimum. When IPIs were eliminated, subjects failed to show a significant orienting bias to the grunt and shrill bark, whereas they continued to turn right in response to the copulation scream. When IPIs were stretched to the maximum, no orienting bias was observed in response to grunts and shrill barks, though many subjects now turned to

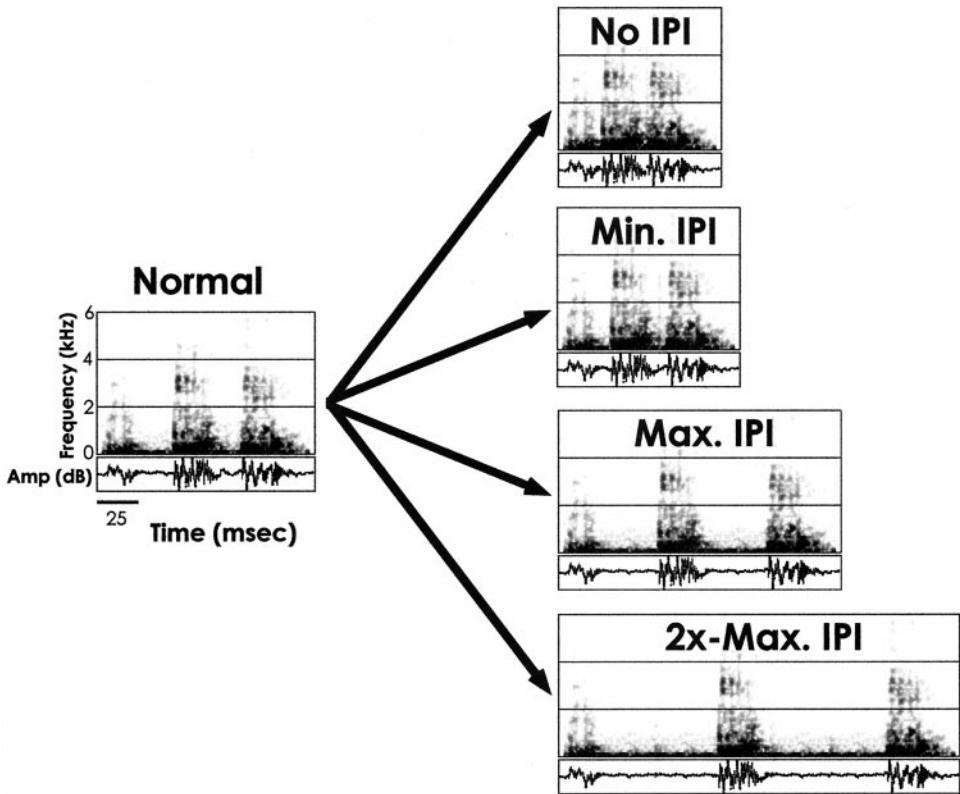


Figure 2. Experimental design for playback experiment on orienting biases in rhesus monkey perception, with a focus on the acoustic morphology of species-specific vocalizations. Shown on the left is a spectrogram and time-amplitude waveform of a normal, unmanipulated 'grunt' given during affiliative interactions. On the right are four manipulations, involving the elimination of interpulse interval (No IPI), reduction of the IPI to the minimum in the population, stretching IPI to the maximum in the population, and stretching it to twice the maximum in the population.

the left or failed to respond; subjects continued to turn right in response to the copulation scream. Lastly, when IPIs were stretched to twice the maximum in the population, subjects consistently turned left in response to the grunt and shrill bark, but maintained a right ear bias for the copulation scream.

These results show that interpulse interval is a salient feature for classifying grunts and shrill barks as rhesus monkey calls, assuming that the orienting bias we have observed is a satisfactory diagnostic (see discussion in Hauser, Agnetta, & Perez, 1998). In contrast, manipulating this feature fails to influence the orienting response to copulation screams. Interestingly, whereas grunts and shrill barks are always produced with a minimum of two pulses per call, copulation screams are often produced with a single pulse. Consequently, although the number of pulses in a call may be correlated with some other property of the caller, such as his genetic quality (e.g., Hauser, 1993b), it is not relevant to this particular call's characteristic morphology.

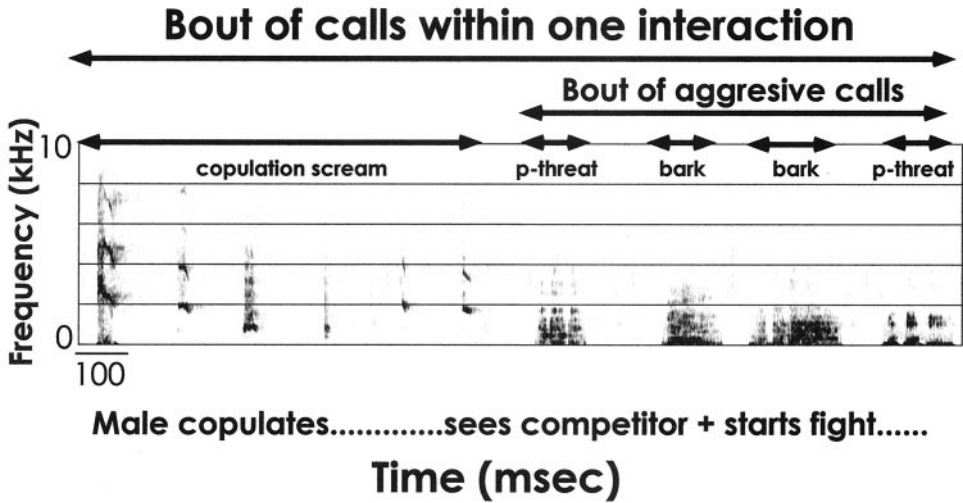


Figure 3. Sound spectrograms of a bout of calls produced by a male rhesus monkey during copulation and then during an aggressive interaction.

Given the internal structure of grunts, shrill barks and copulation screams, it is now possible to carry out more significant manipulations with respect to the unit problem, mixing up the pulses from different call types, or rearranging other acoustical components, to assess if and how subjects respond to such changes. For example, how would subjects respond to a signal that combined a pulse from a grunt with one from a copulation scream? Does the most salient component of the signal dominate, or do such novel combinations generate vocalizations with completely new functions? If a new function emerges, does it represent an averaging of the components (a form of blending) or something entirely novel (the outcome of a particulate system)? At this stage, we simply do not know. And although there are no technical problems in the way (i.e., the signals can readily be manipulated, moving components around as if we were responsible for determining the rules of recombination), we are still confronted with the problem of an appropriate assay, of how to decide whether an individual considers a signal with pulses from a grunt and shrill bark as having new functions or not. We will return to this problem.

Stringing Units Together

Many of the utterances produced by primates occur in bouts or strings of calls (for acoustic analyses of bouts, see Todt, Hammerschmidt, Ansorge & Fischer, 1995). For example, Figure 3 provides an example of a call bout involving a male rhesus monkey who begins the interaction by copulating with a female and uttering a copulation scream. As he finishes mating, he sees a male competitor, dismounts, and begins a bout of aggressive calls, starting with a mild threat ('pant-threat'), moving to a more intense threat ('barks') and then terminating the sequence with a final pant threat as the competitor runs away. We can say that each call functions as a unit in that animals produce these signals alone and when listeners hear them, either naturally or through playbacks, they produce

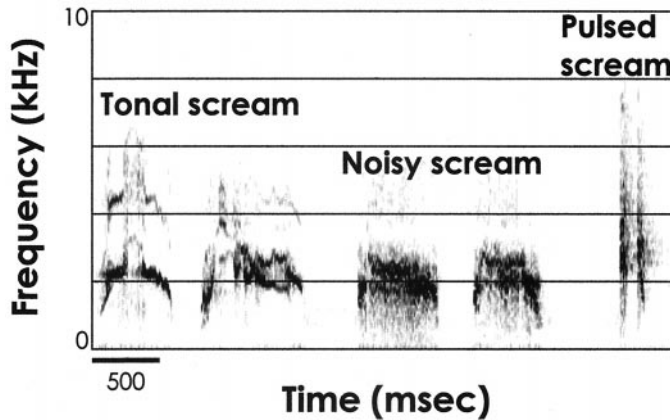


Figure 4. Sound spectrograms of three different scream types produced by a subordinate during an interaction with a dominant.

behaviorally appropriate responses, at least in a general fashion. For example, a dominant's bark can cause a subordinate to back away and a subordinate's scream can halt a dominant's attack. Mixed call bouts are also observed during more narrowly defined contexts, such as when subordinates scream while being attacked by a dominant. Although early work by Gouzoules, Gouzoules & Marler (1984) suggested that rhesus monkeys produce "pure" bouts, with only one scream type per string of calls, more recent but preliminary work suggests that individuals often produce heterogeneous bouts, with different scream types and a characteristic order of call types within the sequence (Hauser, 1999). For example, when an individual sees an approaching dominant, they often start with harmonically structured 'tonal screams'. If the dominant animal continues the approach and then attacks the subordinate, the victim changes to loud, broad band 'noisy screams'. As the dominant leaves, the subordinate shifts again, producing brief, repeated 'pulsed screams'. This order is common (see Figure 4), even though individuals certainly can, and do sometimes produce only one call type over and over again within a bout. Of interest in these cases, then, is the extent to which there are strict constraints on how different call types are ordered within a bout, and whether there are acoustic cues that indicate the start and end points of a bout, revealing coherence at a higher level than the call. I turn now to this second possibility.

When humans speak, independently of the language in which they are speaking, the pitch of the voice typically starts high at the beginning of the sentence and then declines. A proportion of this change in frequency can be accounted for by physiological factors. Specifically, as the lungs deflate during sentence production, the vocal folds vibrate more slowly, the laryngeal muscles tend to relax, and consequently, the fundamental frequency of the utterance drops. Nonetheless, some studies indicate that speakers exaggerate the starting and ending frequency of the utterance, and that such changes provide important information about when a phrase will end. In essence, fundamental frequency declination represents a universal paralinguistic feature, one that appears to mark phrasal boundaries

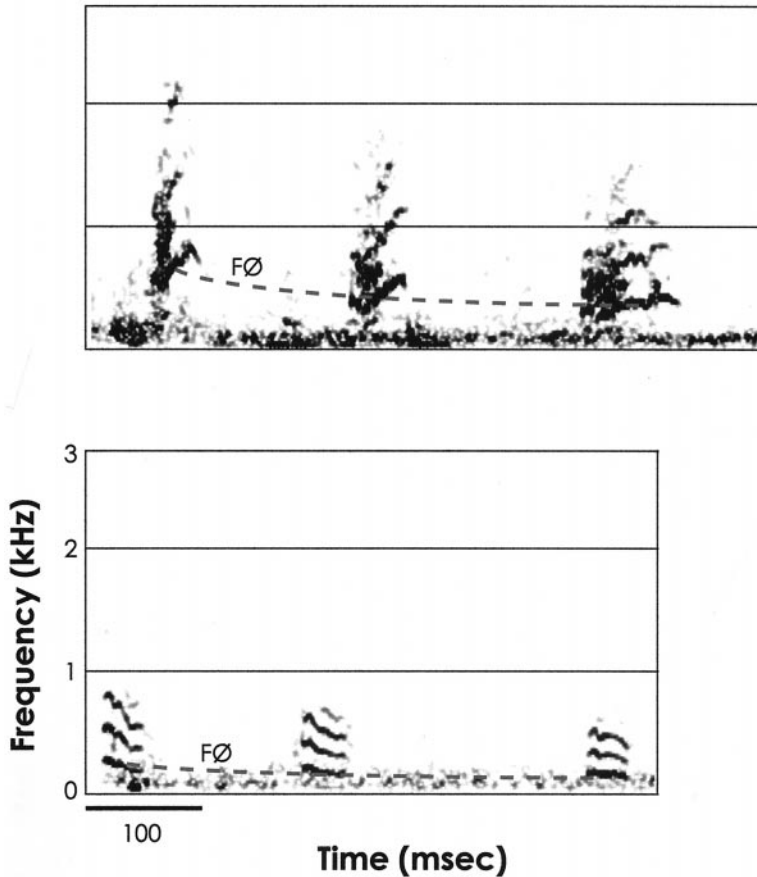


Figure 5. Spectrograms of two different bouts consisting of the rhesus monkeys' 'girney' vocalization. The dashed curve shows the decrease in the fundamental frequency over the course of the bout.

(Gelfer, Harris & Baer, 1987; Pierrehumbert, 1979). Given that the physiological factors underlying this pattern are shared between humans and animals, one might expect some animals to show similar effects. In particular, for those species that produce vocalizations during expiration and where the length of an utterance is limited by the volume of air that can be maintained at any point in time in the lungs, we should expect to see declination in the fundamental frequency. To date, this prediction has been supported by acoustic analyses of two nonhuman primate species, the East African vervet monkey and the rhesus macaque (Hauser & Fowler, 1992). I focus here on the rhesus data.

In rhesus, when individuals engage in close, affiliative social interactions, they often produce a vocalization called the 'girney'. The most common context for production is when mothers approach each other and start inspecting their offspring. Here, the girney is volleyed back and forth in what appears to be a conversational fashion. Consequently, rather than producing a single call, they produce a bout of such calls. Figure 5 presents

sound spectrograms of two girney bouts produced by two different females. What is clear from these two examples, and supported by analyses of several other bouts, is that the average fundamental frequency is highest at the start of a bout and then drops, though the rate at which frequency falls varies from bout to bout. Analysis of the social context associated with these acoustic signals indicates that most bouts are produced in the absence of another animal calling. In other words, when someone calls, they complete their utterances without interruption. When a bout is interrupted by another caller, however, the structure of the bout changes. Rather than a descending frequency contour, there is an increase in the frequency of the call immediately following the interruption. This deviation from the characteristic frequency declination was rarely observed for bouts that were not interrupted. Together, these data provide circumstantial evidence that frequency declination is a perceptually salient acoustic parameter and may be used as an acoustic marker for determining the end of a calling sequence. Further work is needed, using synthesized signals to alter the frequency contour of a bout and determine whether such changes are perceptually salient and biologically meaningful in that they influence the pattern of vocal behavior among interactants.

To summarize this section, studies of primate vocal behavior typically assume that the call is the fundamental unit of analysis. Support for this assumption comes from the fact that calls play a functional role in the communicative system of several primate species. However, by focusing almost exclusively on the call, we are left with a great deal of uncertainty about other possible levels within the system. Understanding whether components of a call can be recombined with others to create new, functionally significant utterances, is a critical first step in addressing the challenges of linguists who claim that animal repertoires lack syntactic structure (Bickerton, 1990; Chomsky, 1986; reviewed in Hauser, 1996; Lieberman, 1984; Pinker, 1994; Studdert-Kennedy, 1998). If there are units that can be recombined, we then need to establish the extent to which the observed variation maps onto variation in emotional expression, referential expression, or some combination of the two. I now turn to the possibility of referential signals.

III. WAYS OF MEANING

“... meanings ain’t in the head ...” (Putnam, 1975)

Historically, the common view of animal communication was that each vocal utterance corresponded to a particular emotion; screams for fear, barks for aggression, and coos for comfort (Smith, 1977; reviewed in Cheney & Seyfarth, 1990a; Hauser, 1996). Seyfarth, Cheney, and Marler’s work on vervet monkey alarm calls raised the possibility that at least some animal sounds were *also* referential, by which they meant that a listener could derive relevant information about the context (e.g., target object or event) eliciting the call from its acoustic properties alone. In this sense, so they argued, vervet alarm calls were like some human words, picking out particular referents *and* indicating the emotional state of the speaker. What was less than clear, and explicitly acknowledged by these researchers, was the precise referent of the call, and whether the caller vocalized with the explicit intent

to refer to a particular object or event. It is these conceptual issues that I would like to focus on, and in particular, the kinds of experiments on primates that Cheney, Seyfarth, and others have conducted to refine the level of analysis and thus, understanding. Due to the focus on nonhuman primates in this paper, I am leaving out comparably elegant experiments on domestic chickens (Evans & Marler, 1995; Marler, Duffy & Pickert, 1986a,b), as well as work on nonhuman primate gestures (Tomasello, Call, Nageli, Olguin & Carpenter, 1994) and child language acquisition (Baldwin, 1995; Bloom, 1994; Bloom & Markson, 1998; Carpenter, Nageli & Tomasello, 1998; Markman, 1990; Pinker, 1994). Due to the conceptual focus, I am also leaving out an important entry into the dictionary, the rich emotional expressions of primates and the functional role they play in mediating social interactions (reviewed in Hauser, 1996; Owren & Rendall, 1998).

Let me start with four assumptions. First, I assume that the call represents the lowest level in terms of functional units within the repertoire. Although I suggested in the first section of this essay that we should withhold judgment on this problem, at present we lack evidence that primates take elements of existing calls and recombine them into calls with new meaning. If this assumption is correct, then we need only worry about the meaning of single calls and when produced, strings or bouts of calls.

Second, although the actual sound used in a particular situation is arbitrary (e.g., the alarm call given to a leopard doesn't sound like a leopard), let us assume that sound-meaning pairs are stimulus bound and that animals lack the capacity to vocalize about possible situations or previously encountered ones; animals clearly have a representation of time (reviewed in Gallistel, 1990; Shettleworth, 1998), but to our knowledge, their vocalizations apparently do not capture the distinction between past and future events. In this sense, animal vocalizations are significantly different from many human words. We can talk about the dangers of encountering a leopard, or what it was like when we ran up a tree in response to being chased by a leopard. Vervets, in contrast, only 'talk' about the leopard they currently see, either out on the plains or hidden behind a bush; they apparently do not warn others of potential dangers or communicate about prior encounters.

Third, I assume that animal vocalizations only refer to whole objects (leopard) or events, and not object fragments (undetached leopard parts) or components of an event. Here then, we are assuming that animal vocalizations are like the words used by children during the early stages of development, a phase documented in detail by Ellen Markman (1990, 1984; see also Bloom & Markson, 1998; Soja, Carey & Spelke, 1991). In particular, early on in development, the child assumes that unfamiliar words refer to whole objects, not parts of an object or objects and the scenes to which they belong. When at the zoo, a child hearing the word 'leopard' doesn't assume it refers to their paws, tail, or spotted fur, and neither does she assume it refers to the leopard and its cage, or the leopard and an observant crowd. Similarly, when vervets produce leopard alarm calls, it seems unlikely that they are referring to its paws or tail. In fact, from an evolutionary perspective, it is difficult to imagine a pressure that would have selected for such fine-grained, part/whole distinctions. By assuming that primates lack this distinction, we can bypass some of Quine's indeterminacy problems. 'Undetached leopard part' is simply not on the list of potential meanings for vervet monkeys.

Fourth, I assume that we can study the referential properties of animal vocalizations independently of the intentions of the caller. For some researchers studying child language acquisition, to have an understanding of referentiality is to have an understanding of the speaker's intentions to refer (see, for example, Baldwin, 1995; Carpenter et al. 1998). Although a complete analysis of animal vocal communication must ultimately combine work on reference and intention, I believe, as do other ethologists (e.g., Cheney & Seyfarth, 1990a; Marler, 1985; Marler, Evans & Hauser, 1992), that we can make significant progress by starting with the problem of reference, and assuming that the caller is attempting to provide information to potential listeners. Although animals may lack a full blown theory of mind (Cheney & Seyfarth, 1990a,b; Premack, 1986; Tomasello & Call, 1997), and thus lack a rich understanding of intentionality, they are not acting reflexively or involuntarily. Several studies show that whether or not an animal calls is mediated by the social audience (i.e., the potential pool of perceivers). Thus, a rooster is more likely to give food calls when a hen is nearby than when another rooster is in view, and vervet females are more likely to give alarm calls to predators when they are with kin than with nonkin. These audience effects show that at least some animal vocalizations are produced with the goal of altering another animal's state. They are not automatic reflexes, elicited by a singular event in the environment.

The Classic Case of Vervet Monkey Alarm Calls

For the sake of discussion, consider one piece of the vervet alarm call story, their alarm call to birds of prey or what Cheney and Seyfarth have referred to as the 'eagle alarm call'. There are two critical points. First, when birds of prey are detected, vervets respond by scanning up to the sky. If impending danger is perceived (the bird is swooping), they head for cover, moving swiftly into dense vegetation. Second, when vervets hear an alarm call given to a bird of prey, they immediately look up and scan the sky; sometimes they head for dense vegetation without scanning. Taking points one and two together, and adding on the vervets' response to other predators (e.g., running up into a tree for leopards, standing bipedally for snakes), we obtain the following uncontroversial conclusion: the eagle alarm call, but no other call, causes the vervets to scan for something in the sky (as opposed to other possible target areas) and to head for dense vegetation (as opposed to other potential places in the environment).

The vervets' highly specific response to birds of prey, and the acoustically distinctive call that accompanies such encounters, raise important questions about the kinds of representation that underlie their behavior. For example, upon hearing an eagle alarm call, what are they looking for when they scan the sky? Are they looking for something that matches a stored representation of a dangerous predator and if so, what is the content of their representation? We know from developmental studies (Seyfarth & Cheney, 1980, 1986) that vervet infants give eagle-sounding-alarm calls to birds that fall outside the class of predators (e.g., vultures) and that, with time, this class narrows. Consequently, it appears that vervets are born with an innate representation of something like 'dangerous things in the sky'. With time, experience weeds out the inappropriate items and selects for

the appropriate ones. But even with this developmental story in hand, we are not yet ready to identify the content of the adult's representation. A second set of experiments, however, firms up the general distinction between predator types, and thus, the general referential distinctions.

An adult vervet monkey A is giving eagle alarm calls while sitting in a bush. Following the first alarm call, listeners respond intensely, looking toward the caller and up toward the sky. With consecutive repetitions of the eagle alarm call from A, and no eagle, however, listeners begin to habituate. If A now begins to produce leopard alarm calls, will listeners perk up, noticing the difference? And if a difference is noticed, is it based on the acoustics or on the referents? Will listeners perk up if vervet B starts producing eagle alarm calls? Here, the type of alarm call hasn't changed, but the identity of the caller has, and we know from previous work that each individual has a distinctive voice (Cheney & Seyfarth, 1980). Cheney and Seyfarth (1988, 1990a) simulated this scenario with playback experiments. Results showed that when either the call type or the caller changed following habituation, subjects dishabituated, increasing the amount of time looking toward the speaker; during the switch to leopard alarm calls, the nature of the response also changed from scanning up to moving toward and up into trees. Clearly, the vervets noticed a difference in the identity of the caller, and the functional difference between eagle and leopard alarm calls. Here again, therefore, we have evidence that eagle and leopard alarm calls designate different things, but our understanding of the underlying representation is still quite poor. Furthermore, studies of alarm calls, both in vervet monkeys and other species (Macedonia, 1991; Zuberbuhler, Noe & Seyfarth, 1997), give us a false sense of security with respect to the prospects of understanding the content of an animal's representation. Unlike the distinctive responses elicited by alarm calls, most primate vocalizations, like most human words, fail to elicit overt, call-specific responses. Rather, most calls elicit an orienting response toward the caller. And yet, some of the more interesting vocalizations produced by primates fall into this class, leaving us with a difficult methodological challenge with regard to the problem of translation. Let me illustrate with an example.

Complications from Other Species and Systems

Rhesus monkeys on Cayo Santiago produce five acoustically distinctive vocalizations when they find food. Three vocalizations ('warble', 'harmonic arch', 'chirp') are restricted to the discovery of high quality, rare food items whereas the other two ('grunt', 'coo') are given while waiting for and eating lower quality, common food items. An individual's hunger level is positively correlated with call rate, but not with the type of call produced. That is, the type of vocalization produced is influenced by the type of food discovered, and not by the discoverer's hunger level (Hauser, 1999; Hauser & Marler, 1993a).

To determine how rhesus classify food-associated calls, a habituation-dishabituation experiment (Hauser, 1998b) was conducted using two of the vocalizations given to high quality/rare foods (i.e., warble and harmonic arch) and one vocalization given to lower quality/common foods (i.e., grunt; see Figure 6). Like Cheney and Seyfarth's experiments

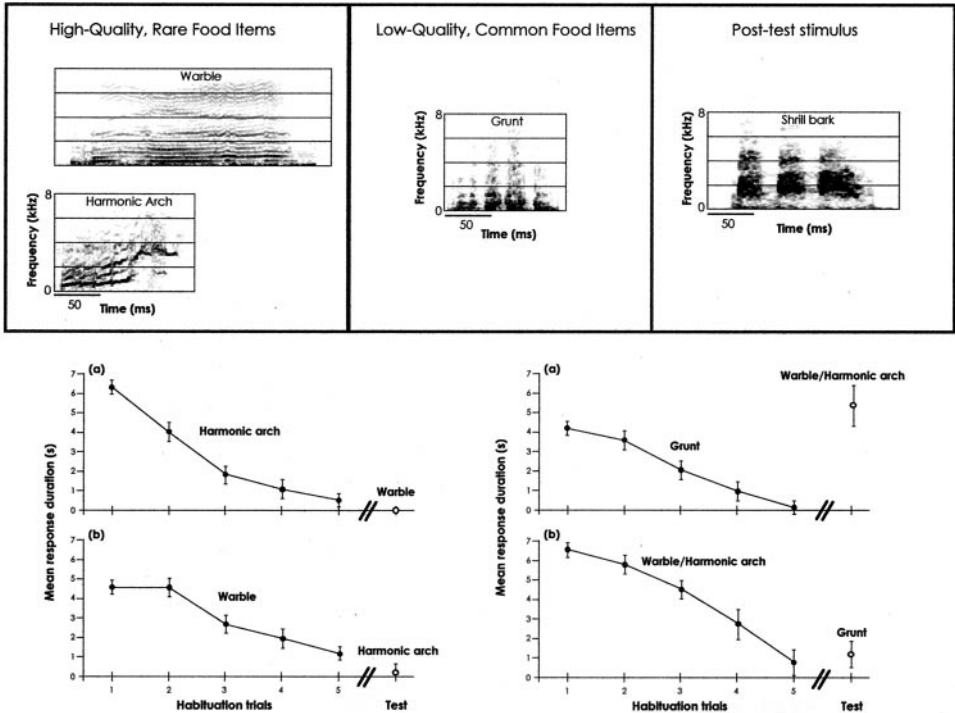


Figure 6. The upper portion of this figure shows sound spectrograms of two food-associated calls ('warble', 'harmonic arch') given to high quality/rare food items, one food-associated call ('grunt') given to low quality/common food items, and an alarm call ('shrill bark') used as a post-test stimulus in the playback experiments. Experiments were run by habituating individuals to exemplars from one class of stimuli and then, following two consecutive trials with no response (i.e., no orientation to the speaker), testing them with one exemplar from a different class. If subjects failed to respond on the test trial, a post-test was run to insure that they had not simply habituated to all sounds emanating from the test area. The lower portion of the figure shows the results from the playback experiments, with the mean amount of time spent looking toward the speaker (seconds) on the y-axis and the number of trials on the x-axis.

on vervet alarm calls, our playbacks were designed to distinguish between two competing hypotheses. Specifically, do rhesus classify vocalizations on the basis of acoustic features alone or on the basis of a call's referent? Two conditions were run. In the first, we contrasted the warble and harmonic arch. Following habituation to warbles, would rhesus transfer habituation or dishabituate? Transfer of habituation would indicate that despite acoustic differences between these two call types, subjects perceptually clustered them into one category based on something like similarity of the putative referent (i.e., high quality/rare food); to test for perceptual symmetry in this condition, we also habituated subjects to harmonic arches and then tested them with a warble. Dishabituation would either indicate that perceptual classification is based on acoustic differences or, that there are significant differences in the referents of warbles and harmonic arches, differences that our observations failed to reveal. The second condition involved habituation to either the

warble or harmonic arch followed by a test with the grunt; here too we were interested in the possibility of perceptual symmetry and thus habituated some animals to grunts, and then tested them with either a warble or harmonic arch. Transfer of habituation would indicate perceptual clustering on the basis of shared referents; as in the first condition, dishabituation could be due to both acoustic and referential differences. By using each call type as both a habituating stimulus and as a test stimulus, our experiment would allow us to examine asymmetries in the magnitude of the subject's response in the test trial. Should asymmetries emerge, there would be stronger support for the referential hypothesis, since the stimuli compared are equivalent within each condition. For example, the acoustic differences between a grunt and a warble are held constant, regardless of which call serves as the habituating stimulus and which as the test stimulus; differences in response are therefore likely to be due to referential differences between the habituating and test stimuli.

Results (Figure 6) showed that following habituation to a warble, subjects consistently transferred the level of habituation to harmonic arches. Similarly, following habituation to harmonic arches, subjects transferred habituation to warbles. Thus, although these two call types are acoustically and perceptually discriminable, the rhesus appear to cluster them into one category. This suggests that warbles and harmonic arches share a common referent. In contrast, following habituation to either warbles or harmonic arches, subjects consistently dishabituated to grunts, and the same pattern emerged when the habituating and test stimuli were reversed. However, the magnitude of the dishabituating response to grunts was significantly less than it was to either warbles or harmonic arches. This suggests that subjects perceive a difference between grunts and warbles/harmonic arches, a difference that lies primarily in their referents rather than their acoustics.

Consider, as an analogy to the rhesus case, a translation of warble into 'caviar', harmonic arch into 'fish eggs' and grunt into 'potato'. Now imagine that you hear someone at a restaurant repeatedly saying 'potato' as he eats forkful after forkful. Eventually, you habituate and carry on eating your own meal. If, all of a sudden, your vocal neighbor says 'caviar', you would presumably perk up and orient, not because of an acoustic difference but because of a referential difference, and a meaningful one at that. Conversely, if he starts by repeating 'caviar' and then switches to 'fish eggs', an acoustic difference would certainly be noticed, but there is only a trivially important referential difference. Presumably, you would keep eating, ignoring this change. My guess is that the rhesus data can be similarly interpreted. We are still left, however, with an unsatisfactory understanding of the content of their representation when they produce such calls, as well as when they hear them. Do warbles refer to the same things as harmonic arches, or are their subtle differences, covarying with the type of rare, high quality food discovered? To date, there is no evidence that warbles are given to one type of food and harmonic arches to another. This may, however, reflect the limited variety of food types on the island of Cayo Santiago, and the fact that a majority of the diet consists of provisioned monkey chow (Hauser, 1999).

Summarizing, primate vocalizations can be more than mere expressions of the caller's emotional state. They can be functionally referential, conveying information about salient objects and events in the environment, allowing individuals to make adaptive responses

even when the context associated with call production is not immediately visible (i.e., the listener is out of sight of the caller). Although the nature of the representation evoked by such calls has yet to be precisely specified, studies of alarm calls provide some of the most significant insights (Cheney & Seyfarth, 1990a; Macedonia, 1991; Zuberbuhler, Noe & Seyfarth, 1997; reviewed in Hauser, 1996). When an animal hears an alarm call and begins scanning in a particular location, they must be searching for something specific. The call does not merely cause individuals to engage in a specific, stereotyped motor response. If it did, we would have no evidence of referential information. Rather, we would have a situation that is analogous to a rat pressing a button in response to hearing a tone in a Skinner box. But because the animal scans after hearing an alarm call, we are presented with more than a stereotyped motor response. The animal scans for a target object or event, requiring the capacity for object recognition and for matching a target with an expectation. What is missing from our analysis, however, is whether the expected representation is of a specific kind of predator (e.g., a leopard, a martial eagle) or something much less rigid, but functionally more flexible and adaptive, something like a prototype for the general category (e.g., large cat-like, bird of prey-like). Although it is clear that animals discriminate predator kinds (the different alarm calls and escape responses provide unambiguous evidence for this), we know little about their representation of natural kinds. In this sense, studies of vocal communication must attach themselves in a more significant way to studies of conceptual representation, and especially, investigations into how animals represent different kinds of objects, including artifactual, animate, inanimate, and so on (Cheney & Seyfarth, 1990b; Cheney, Seyfarth & Silk, 1995; Hauser, 1997, 1998a; Premack, 1986; Thompson, 1995).

IV. ESTABLISHING CREDIBILITY

“Roughly, if you start out with a true thought, and you proceed to do some thinking, it is very often the case that the thoughts that the thinking leads you to will also be true. This is, in my view, the most important fact we know about minds; no doubt it’s why God bothered to give us any. A psychology that can’t make sense of such facts as that mental processes are typically truth preserving is ipso facto dead in the water.” (Fodor, 1994; p. 9).

Thought requires content. Content is typically acquired by having experiences, be they acquired by an individual over the course of its life, or acquired by the species and innately encoded in the brain; the latter evolve when experiences correspond to statistical regularities of the environment, ones that all individuals confront. To have a *leopard* thought is to think about leopards in a way that is, to some approximation, like experiencing a leopard or something like one, seeing him run across the savanna, hearing him give a territorial roar, uncovering his tracks in the sand, being told that a leopard is like a fat cheetah and a thin lion, and so forth. By holding a particular relationship to the world, thoughts carry a certain truth value. I hear a growling sound and think that it sounds like a leopard. When I move toward the bush where I heard the sound, I find a speaker, not a leopard. My expectation is proved false, and the content of my thought changes. In this section, I turn to the relationship between what a call potentially refers to and how animals

might assess its truth value. My concern with truth is, however, far more modest than it appears. In particular, we will explore two particular cases that will hopefully shed some light on the deeper philosophical problems. The first focuses on honesty, and how particular features of a communication system constrain the extent to which one animal can functionally deceive another. This problem can be approached without delving into the conceptually more challenging issues of intentionality and beliefs, problems that other authors in this volume address. The second case also bypasses such issues, focusing instead on an animal's expectations, on its capacity to detect violations between a vocal utterance and an anticipated outcome.

Honesty, Credibility, and Economics

During the early days of ethology, studies explored the origins of communicative signals, of how noncommunicative actions developed into stereotyped and ritualized signals for conveying information (Hinde, 1981; Lorenz, 1966; for example, the honey bees' dance language was thought to have originated from a flight intention movement associated with departure to a foraging area (Gould & Gould, 1988; von Frisch, 1967). In the mid-1970s, when sociobiological perspectives emerged, focusing on selfish genes and manipulative behaviors, the ethological position on communication was challenged (Dawkins & Krebs, 1978; Maynard Smith, 1974). In particular, in games of competition, why would one animal show its cards, providing its competitor with information about the next move (e.g., flee)? Such competitive situations are vulnerable to cheaters, individuals who continuously escalate the level of aggressive intent, but in reality, have no way of defending their position. An emerging view, then, was that animals were designed to produce signals that manipulated the behavior of recipients, holding a poker face along the way. Receivers, in contrast, would be designed to mind-read, to detect falsehoods, and find truth in advertising (Grafen, 1990; Krebs & Dawkins, 1984; Zahavi, 1975). But how? Over the past twenty years, there has been an explosion of studies on honest signaling in animals, with research focused on the extent to which animals are capable of lying about the current state of affairs and the extent to which lies can be detected (reviewed in Hauser, 1996; Zahavi & Zahavi, 1997). These studies are absolutely silent with respect to the cognitive or conceptual tools needed. Instead, they focus on the problem from a functional perspective, one interested in the relationship between signal design and reproductive outcome.

To provide a flavor of this argument, consider the following. A subordinate baboon is approached by a more dominant group member. The dominant yawns, an aggressive display that provides a clear view of its canines. The subordinate holds his ground. The dominant escalates, standing bipedally, hair bristling, and barking. The subordinate is now in the position of trying to work out the honesty of the dominant's display, including both the visual and auditory components. This entails making a distinction between two kinds of signaling parameters. First, some morphological features cannot be as readily faked. They are part of the animal's phenotype, part of the equipment that comes with investing in normal growth and development. Thus, the size of a male's canines cannot be faked,

though they can be concealed by yawning with pursed lips, or by not yawning at all. When the canines are revealed, they provide an honest indication of the male's fighting power. When the male barks, the formant frequencies that characterize the call reflect the length of the caller's vocal tract. And as Fitch (1997; Fitch & Hauser, 1995) has demonstrated through anatomical and acoustical analyses of several primate and nonprimate species, the length of the vocal tract may anchor honesty during vocal exchange. Unlike the larynx, there are greater physical constraints on adjusting vocal tract length, constraints that are imposed by tight correlations with body size and skull size; for example, though an animal can make itself sound smaller or larger by retracting or protruding the lips respectively, the change imposed is highly constrained. As a result, animals attending to the formants of a call can readily extract reliable information about the caller's size.

Body size, like canine size, is also difficult to fake. There is, however, an interesting twist to this story given the capacity of some mammals to piloerect. By bristling their hair, individuals give the appearance of a bigger animal; there are many comparable examples in the animal kingdom, including the pneumatic increase in size by blowfish, the fanning of feathers around the head or tail in some birds, and the hood displays of frilled lizards and spitting cobra. These examples bring up the second signaling parameter, in particular, the ability of perceivers to assess the relationship between a display and the displayer's capacity to back up the apparent message in the display. When the dominant signals aggressive intent by piloerecting and barking, is he capable of following through with the message conveyed by the display, chasing and physically attacking the subordinate? If not, then his display represents a bluff, a dishonest signal of intent or motivational state. It is this second parameter that is of interest to us here.

The evolutionary biologist, Amotz Zahavi (1975, 1993), provided one answer to this problem. Specifically, he suggested that under some conditions, honesty could be anchored in the relative costs associated with producing a signal. For example, a dominant male who piloerects and barks, but shows signs of weakness (e.g., ribs showing from malnutrition), would, on this account, be considered dishonest. If the verity of the dominant's display is challenged, presumably he would back off; the benefits of continuing the aggressive bluff would be outweighed by the costs of a subordinate who turned around and attacked. Studies of nonprimates in particular have found some evidence that costs are related to signaling honesty (FitzGibbon & Fanshawe, 1988; Hauser, 1993b; Møller, 1993).

Communication and Expectation

In discussing studies of referential signaling, and in particular, whether animals compare vocalizations on the basis of their acoustic properties alone or on the basis of their putative referents, we actually touched upon the problem of signal honesty. If an audio speaker repeatedly transmits individual A's eagle alarm call or food-associated call, subjects eventually habituate. Habituation could be interpreted as evidence that listeners consider A to be unreliable when it comes to providing *veridical* information about eagles or food. However, if you now play the same call from a different group member, B, subjects

respond. Although A may be unreliable about the presence of an eagle, B need not be. This is precisely what Cheney and Seyfarth (1990a) demonstrated for vervet monkey alarm calls and intergroup calls. Let me unpack this a bit.

When an individual hears an alarm call or food call for the first time, its orienting response is presumably guided by an interest in confirming or disconfirming some kind of expectation, of finding out whether an eagle or some food has, in fact, been discovered. During playback experiments, the target referent is never present. Thus, with the delivery of each playback in the habituation series, listeners gain increasing confidence that the mapping between call and world, or between referent and expected representation, is false. On a simple level, then, the truth value of an utterance can be derived by a kind of crude statistical analysis. This simple-minded statistical analysis fails, however, when we consider the range of factors that might influence an individual's reliability, including their age, tenure in a group, the number of interactions that have been exchanged in the past, and differences that might arise across contexts (e.g., announcing the discovery of food, the position of a neighboring competitor, the approach path of a predator). In the vervet and rhesus monkey experiments, then, such issues could be addressed by measuring the rate of habituation as well as changes in the patterns of habituation over the course of several days. For example, if you habituate an individual to A's eagle alarm call on day one, how do they respond to a second habituation series on day two? After repeating this habituation series over several days, what would it take to reinstate A's reliability? Would one correct call—a hit in the signal detection theory sense—be sufficient? Is A's reliability with respect to eagles at all influenced by his reliability in other contexts? We don't have answers to these questions, but the experiments are relatively easy to run. Let me end with two examples that may help illuminate the connection between meaning and honesty.

In birds and mammals, young produce distress calls or cries that typically function to recruit care from adults, most often the offspring's mother. In vervet monkeys, infants give distress calls when attempting to nurse and when attempting to ride on their mother during travel (Struhsaker, 1967). Early on in development, mother always respond to their infant's requests (Hauser, 1993a). At the start of weaning (approximately 3 months), infants escalate the rate of distress call production while mothers increase the frequency with which they reject their infants requests for care (e.g., attempts to nurse). Some infants appear to perceive this change in call effectiveness, and drop the rate of calling. Apparently in response to this change, call effectiveness rebounds, although not to the level achieved prior to weaning. Other infants continue calling at high rates and effectiveness continues to plummet. Many of these infants die within a few months. One interpretation of these data are that call rate provides a dishonest indication of infant needs. Infants that call beyond their needs (i.e., those required to support normal growth, development and survival) are ignored, and suffer the consequences. In contrast, infants that adjust call rate to match physiological needs, obtain the level of care requested. An alternative interpretation, one that cannot be ruled out, is that mothers differ with respect to the level of care they are capable of providing. Mothers in poor condition are unable to respond to their infant's requests, even though their distress calls may provide veridical indications of need. Weighing against this alternative view, however, is the observation that infants who

failed to alter call rate, and subsequently died, were born in both high and low quality territories, and to mothers who were both high and low ranking. Thus, differential access to high quality resources, which can be considered a proxy for condition, did not explain the pattern observed. More detailed analyses of physical condition, along with vocal and behavioral data, are needed to determine the informational content of such calls, and the degree to which a truth value can be attached to them.

As a final example, let us return to rhesus monkey food calls (Hauser & Marler, 1993a,b). When individuals find food, several factors appear to influence whether or not they call, the kind of call they give, and the rate at which they call. In particular, based on natural observations and field experiments, females call more often than males, both sexes call more often when they are hungry than when they are satiated, and the type of call produced is determined by the type of food discovered. Under experimental conditions, involving the presentation of food to individuals who are alone and out of visual contact from all other individuals, subjects who fail to call receive more aggression and obtain less food than subjects who call; this effect holds even if the discoverer is a high ranking individual. Here, calling appears to convey the information that food has been discovered, whereas the absence of calling (i.e., withholding information) appears to violate a convention or rule, as evidenced by the observation that silent discoverers are physically attacked. Complicating the story is the observation that the above pattern only holds for members of a social group. When peripheral males (individuals who have yet to join a social group) are targeted as discoverers, they never give food calls and when they are detected by members of a social group, they are never attacked. Can we conclude from these observations that silent discoverers are dishonest? Can we *ever* interpret the absence of a response in terms of the individual's intentions? Maybe. These questions are central to the problem of meaning and honesty because they force us to consider what individuals believe about others, what they know about the function of their calls and the effects they have on others with respect to behavior and beliefs. At present, and as several essays in this issue reveal, we know little about mental state attribution in nonhuman animals. What is known, however, suggests that the ability to infer mental states in others is either absent, or weakly expressed. And without this capacity, assessing whether the expressed meaning of an utterance is honest or dishonest will be based entirely on behavioral encounters, an assessment that is driven by statistical regularities.

Summarizing, studies of meaning and honesty have generally focused on problems of evolutionary function and significance, rather than on problems of underlying mental states. Results suggest that animals are capable of detecting violations, ones based on the relationship between call meaning, an individual's expectations, and observed outcomes or consequences. To flesh out the nature of such expectations, additional experiments are necessary. Fortunately, some of the tools are in place.

V. METHOD AND MEANING

In this essay I have developed three points concerning the problem of decoding, of converting the vocal utterances of a nonhuman species into something that we, as

non-native speakers, can interpret. Specifically, I have tried to argue that progress in this area will depend on refining our understanding of functional units within the repertoire, the putative referents of a signal, and the conceptual representations and intentional states that underlie the production and perception of vocalizations. To conclude, I focus on a few methodological issues that I believe have important implications for the conceptual issues at hand.

Central to our discussion of animal communication is the extent to which vocalizations provide representational information about objects and events in the world, and the extent to which they invoke in listeners representations that accurately depict the current state of affairs. If vocalizations invoke specific sorts of representations, then upon hearing a call, individuals will set up expectations about the kinds of events that are likely to transpire in the near future. In this sense, a key to understanding the nature of animal representation is to create situations that are either consistent or inconsistent with their expectations. Consider a simple anecdotal observation, one that can function as an intuition pump for designing experiments.

While working at Cheney and Seyfarth's field site in Amboseli, Kenya, I observed a 6-month-old vervet monkey produce a leopard alarm call to a herd of elephants moving nearby; elephants never prey on vervets, and adults never give alarm calls to them. At almost the same time, an adult male in the infant's group detected a leopard and produced a leopard alarm call. For several months, the infant continued to produce leopard alarm calls to elephants, even though the adults never responded by fleeing or producing additional alarm calls. Given that the infant produced leopard alarm calls to elephants, what can be said about his representation, about his beliefs, his expectations? First, it seems unlikely that vervet infants are incapable of discriminating leopards from elephants. Their visual system is sufficiently well developed at this age to make the distinction. A more likely explanation is that the rapidly moving elephants elicited fear which, in turn, elicited an alarm call. The leopard alarm call is the one most likely to be elicited since it appears to emerge, early in development, in response to large ground predators (i.e., not snakes and not birds of prey). Given the repeated association between elephants and leopard alarm calls over the next few months, what can be said about the infant's expectations? What did it expect to see upon hearing a leopard alarm call? Would it be surprised to see a gazelle, giraffe or hippopotamus, animals that fall outside the category of vervet predators? We don't have answers to these questions, but experiments based on the logic of expectancy violation bring us closer to an answer.

To explore the nature of expectation, and the capacity for monkeys to store information in short term memory, Tinkelpaugh (1928) conducted the following experiment. While a rhesus monkey was seated in a cage, watching, a banana was placed under one of two cups. A screen was then raised, occluding the monkey from the two cups. After a short delay, the screen was removed and the monkey was allowed to select a cup. The cup with the banana was consistently selected, showing that they had little difficulty remembering the location of the banana even when it was placed out of sight for a short period of time. The same kind of performance was also demonstrated with a piece of lettuce, a less

preferred food item. In a final test, a banana was concealed under one cup and while the screen was in position, occluding the monkey's view, the experimenter removed the banana and replaced it with the lettuce. When the screen was taken away, the monkey selected the cup previously containing the banana, only to find the lettuce. Tinkelpaugh's description suggests that the monkey was apparently annoyed, surprised to find the lettuce when it apparently 'expected' the banana. When the same experiment was run, but a banana was substituted for the previously concealed lettuce, the monkey showed no overt signs of surprise. Although it is difficult to interpret Tinkelpaugh's quite casual observations of the monkey's response, it appears that the monkeys detected a violation. More specifically, it appears that they formed a representation of the kind of object concealed, expected this object to be under the cup, and perceived a violation when the object uncovered differed from the one expected. Recent neurophysiological work suggests that this kind of expectation is established by activation of the prefrontal cortex (Watanabe, 1996), an area known to play an essential role in working memory.

Several experiments have now been conducted on nonhuman primates using the expectancy violation procedure (Hauser, 1998a; Hauser, MacNeilage & Ware, 1996). These experiments show that looking time can be used as a measure of the kinds of representations animals form, and the kinds of expectations that emerge from such representations. Only one experiment to date, however, has used the logic underlying this procedure to explore the relationship between vocal communication and expectation. Cheney, Seyfarth, and Silk (1995) designed a playback experiment to ask whether wild baboons understand the causal relationship between certain kinds of vocalizations and certain kinds of social relationship. In baboons, dominant females often 'grunt' to subordinate mothers when interacting with their infants. In response, subordinate mothers often produce submissive 'barks'. Playback experiments were set up to contrast subjects' responses to causally inconsistent and consistent call sequences. In an inconsistent sequence, subjects heard a subordinate female giving a grunt followed by a higher ranking female giving a bark. In the consistent sequence, the same grunt-bark series was played, followed by a grunt from a female who was dominant to both the first and second callers in the sequence; this vocal interaction is now consistent because the female's bark is temporally associated with the more dominant female's grunt at the end of the sequence. Results showed that subjects looked longer at the causally inconsistent sequence than at the causally consistent sequence. On the basis of these findings, Cheney and colleagues suggest that baboons understand the factors that causally link particular vocalizations up with particular social relationships. These experiments, together with other expectancy violation experiments, are critical for they provide a tool to explore the relationship between expectation, representation and vocal communication.

In conclusion, although we are a far cry from publishing the OPD, the Oxford Primate Dictionary, I hope this essay has helped to identify some of the key problems associated with decoding another species' vocal repertoire. To go beyond the current abridged addition will require a more focused effort on the nature of animal representation and the extent to which vocalizations function as vehicles for expressing such representations.

Acknowledgments: For comments on either subsets of the data presented or early versions of the manuscript, I thank members of the Primate Cognitive Neuroscience Lab at Harvard, D. Baldwin, D. Cheney, T. Fitch, J.G. Greeno, P. Marler, R. Seyfarth, D. Rendall, and M. Studdert-Kennedy. Funding for the work reported on this paper was provided by grants from the National Science Foundation (SBR-93-57976; SBR-9602858) and Harvard University (Provost, Innovation, Anthropology).

REFERENCES

- Abler, W. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures*, 12, 1–13.
- Abler, W. L. (1997). Gene, language, number: The particulate principle in nature. *Evolutionary Theory*, 11, 237–248.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9, 321–324.
- Baldwin, D. (1995). Understanding the link between joint attention and language acquisition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 133–156). Hillsdale: Erlbaum.
- Bickerton, D. (1990). *Species and language*. Chicago: Chicago University Press.
- Bloom, P. (1994). *Language acquisition: Core readings*. Cambridge: MIT Press.
- Bloom, P., & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Science*, 2, 67–73.
- Borden, G. J., & Harris, K. S. (1984). *Speech science primer: Physiology, acoustics, and perception of speech*. Baltimore: Williams & Wilkins.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence. *Monographs of the Society for Research in Child Development*, 63(4), Serial 255, 1–174.
- Catchpole, C. K., & Slater, P. (1995). *Bird song: Biological themes and variations*. Cambridge: Cambridge University Press.
- Cheney, D. L., & Seyfarth, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, 28, 362–367.
- Cheney, D. L., & Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, 36, 477–486.
- Cheney, D. L., & Seyfarth, R. M. (1990a). *How monkeys see the world: Inside the mind of another species*. Chicago: Chicago University Press.
- Cheney, D. L., & Seyfarth, R. M. (1990b). The representation of social relations by monkeys. *Cognition*, 37, 167–196.
- Cheney, D. L., Seyfarth, R. M., & Silk, J. (1995). The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning? *Journal of Comparative Psychology*, 109, 134–141.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Clark, A. P. (1993). Rank differences in the production of vocalizations by wild chimpanzees as a function of social context. *American Journal of Primatology*, 31, 159–180.
- Cleveland, J., & Snowdon, C. T. (1981). The complex vocal repertoire of the adult cotton-top tamarin. *Saguinus oedipus oedipus*. *Zeitschrift für Tierpsychologie*, 58, 231–270.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: evidence from juncture misperception. *Journal of Memory and Language*, 31, 218–236.
- Cynx, J. (1990). Experimental determination of a unit of song production in the zebra finch (*Taeniopygia guttata*). *Journal of Comparative Psychology*, 104, 3–10.
- Dawkins, R., & Krebs, J. R. (1978). Animal signals: Information or manipulation. In J.R. Krebs & N.B. Davies (Eds.) *Behavioural ecology* (pp. 282–309). Oxford: Blackwell Scientific Publications.
- Doupe, A., & Konishi, M. (1991). Song-selective auditory circuits in the vocal control system of the Zebra Finch. *Proceedings of the National Academy of Sciences*, 88, 11339–11343.
- Evans, C. S., & Marler, P. (1995). Language and animal communication: Parallels and contrasts. In H. Roitblatt & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 241–282). Cambridge, MA: MIT Press.

- Ficken, M. S., & Popp, J. W. (1992). Syntactical organization of the gargle vocalization of the black-capped chickadee, *Parus atricapillus*. *Ethology*, *91*, 156–168.
- Fischer, J. (1990). Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*, *55*, 799–807.
- Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*, 1213–1222.
- Fitch, W. T., & Hauser, M. D. (1995). Vocal production in nonhuman primates: acoustics, physiology and functional constraints on honest advertisement. *American Journal of Primatology*, *37*, 191–219.
- FitzGibbon, C. D. & Fanshawe, J. W. (1988). Stotting in Thompson's gazelles: an honest signal of condition. *Behavioral Ecology and Sociobiology*, *23*, 69–74.
- Fodor, J. A. (1994). *The elm and the expert*. Cambridge: MIT Press.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge: MIT Press.
- Gardner, R. A., Gardner, B. T., & Van Cantfort, E. (Eds.). (1989). *Teaching sign language to chimpanzees*. Albany: State University of New York Press.
- Gelfer, C. E., Harris, K. S., & Baer, T. (1987). Controlled variables in sentence intonation. In T. Baer, C. Sasaki & K. Harris (Eds.), *Laryngeal function in phonation and respiration* (pp. 422–435). Boston: College-Hill Publishers.
- Gould, J. L., & Gould, C. G. (1988). *The honey bee*. New York: Freeman Press.
- Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: Representational signaling in the recruitment of agonistic aid. *Animal Behaviour*, *32*, 182–193.
- Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, *144*, 475–546.
- Green, S. (1975). Variation of vocal pattern with social situation in the Japanese macaque (*Macaca fuscata*): a field study. In L.A. Rosenblum (Eds.), *Primate behavior* (Vol. 4, pp. 1–102). New York: Academic Press.
- Hailman, J. P., & Ficken, M. S. (1987). Combinatorial animal communication with computable syntax: Chick-a-dee calling qualifies as 'language' by structural linguistics. *Animal Behaviour*, *34*, 1899–1901.
- Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1987). Constraints on the structure of combinatorial "chick-a-dee" calls. *Ethology*, *75*, 62–80.
- Hammerschmidt, K., & Fischer, J. (1997). The vocal repertoire of Barbary macaques: A quantitative analysis of a graded signal system. *Ethology*, *132*: 381–399.
- Harcourt, A. H., Stewart, K. & Hauser, M. D. (1993). The social use of vocalizations by gorillas. I. Social behaviour and vocal repertoire. *Behaviour*, *124*, 89–122.
- Hauser, M. D. (1993a). Do vervet monkey infants cry wolf? *Animal Behaviour*, *45*, 1242–1244.
- Hauser, M. D. (1993b). Rhesus monkey (*Macaca mulatta*) copulation calls: Honest signals for female choice? *Proceedings of the Royal Society, London*, *254*, 93–96.
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge: MIT Press.
- Hauser, M. D. (1997). Artifactual kinds and functional design features: What a primate understands without language. *Cognition*, *64*, 285–308.
- Hauser, M. D. (1998a). Expectations about object motion and destination: Experiments with a nonhuman primate. *Developmental Science*, *1*, 31–38.
- Hauser, M. D. (1998b). Functional referents and acoustic similarity: Field playback experiments with rhesus monkeys. *Animal Behaviour*, *55*, 1647–1658.
- 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.
- Hauser, M. D. (1999). The sound and the fury: Primate vocalizations of thought and emotion. In N. Wahlin & B. Merker (Eds.) *The origins of music* (pp. 77–102). Cambridge, UK: Cambridge University Press.
- Hauser, M. D., & Andersson, K. (1994). Left hemisphere dominance for processing vocalizations in adult, but not infant rhesus monkeys: Field experiments. *Proceedings of the National Academy of Sciences*, *91*, 3946–3948.
- Hauser, M. D., & Fowler, C. (1992). Declination in fundamental frequency is not unique to human speech: evidence from nonhuman primates. *Journal of the Acoustic Society of America*, *91*, 363–369.
- Hauser, M. D., MacNeilage, P., & Ware, M. (1996). Numerical representations in primates. *Proceedings of the National Academy of Sciences*, *93*, 1514–1517.
- Hauser, M. D., & Marler, P. (1993a). Food-associated calls in rhesus macaques (*Macaca mulatta*). I. Socioecological factors influencing call production. *Behavioral Ecology*, *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. *Behavioral Ecology*, *4*, 206–212.

- Heffner, H. E., & Heffner, R. S. (1984). Temporal lobe lesions and perception of species-specific vocalizations by macaques. *Science*, 226, 75–76.
- Herman, L. M., Pack, A. A. & Palmer, M.-S. (1993). Representational and conceptual skills of dolphins. In H. L. Roitblat, L. M. Herman, and P. E. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 403–442). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Hinde, R. A. (1981). Animal signals: Ethological and games-theory approaches are not incompatible. *Animal Behaviour*, 29, 535–542.
- 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.
- Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural Ecology* (pp. 380–402). Sunderland: Sinauer Associates Inc.
- Kroodsma, D. E., & Miller, E. H. (1996). *Ecology and evolution of acoustic communication in birds*. Ithaca, NY: Comstock Publishing.
- Lambrechts, M. M. (1992). Male quality and playback in the great tit. In P. K. McGregor (Ed.), *Playback and studies of animal communication* (pp. 135–152). New York: Plenum Press.
- Lieberman, P. (1984). *The biology and evolution of language*. Cambridge: Harvard University Press.
- Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.
- Lorenz, K. (1966). Evolution of ritualization in the biological and cultural spheres. *Philosophical Transactions of the Royal Society*, 251, 273–284.
- 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.
- Maeda, T., & Masataka, N. (1987). Locale-specific vocal behaviour of the tamarin (*Saguinus l. labiatus*). *Ethology*, 75, 25–30.
- Margoliash, D., & Fortune, E. (1992). Temporal and harmonic combination-sensitive neurons in the Zebra Finch's HVC. *Journal of Neuroscience*, 12, 4309–4326.
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14, 57–77.
- Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, 16, 1–27.
- Marler, P. (1977). The structure of animal communication sounds. In T.H. Bullock (Eds.), *Recognition of complex acoustic signals* (pp. 17–35). Berlin: Springer-Verlag.
- Marler, P. (1985). Representational vocal signals of primates. In B. Hölldobler and M. Lindauer (Eds.), *Experimental behavioral ecology and sociobiology* (pp. 211–221). Stuttgart, Germany: Gustav Fischer Verlag.
- Marler, P. (1997). Three models of song learning: Evidence from behavior. *Journal of Neurobiology*, 33, 1–16.
- Marler, P., Dufty, A., & Pickert, R. (1986a). Vocal communication in the domestic chicken. I. Does a sender communicate information about the quality of a food referent to a receiver? *Animal Behaviour*, 34, 188–193.
- Marler, P., Dufty, A., & Pickert, R. (1986b). Vocal communication in the domestic chicken. II. Is a sender sensitive to the presence and nature of a receiver? *Animal Behaviour*, 34, 194–198.
- Marler, P., Evans, C. S., & Hauser, M. D. (1992). Animal signals? Reference, motivation or both? In H. Papoušek, U. Jürgens, and M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 66–86). Cambridge: Cambridge University Press.
- Masataka, N. (1988). The response of red-chested moustached tamarins to long calls from their natal and alien populations. *Animal Behaviour*, 36, 55–61.
- Masataka, N., & Fujita, K. (1989). Vocal learning of Japanese and rhesus monkeys. *Behaviour*, 109, 191–199.
- May, B., Moody, D. B., & Stebbins, W. C. (1989). Categorical perception of conspecific communication sounds by Japanese macaques. *Macaca fuscata*. *Journal of the Acoustical Society of America*, 85, 837–847.
- Maynard Smith, J. (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology*, 47, 209–221.
- Mitani, J., & Marler, P. (1989). A phonological analysis of male gibbon singing behavior. *Behaviour*, 109, 20–45.
- Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P., & Byrne, R. (1992). Dialects in wild chimpanzees? *American Journal of Primatology*, 27, 233–244.

- Møller, A. P. (1993). Morphology and sexual selection in the barn swallow *Hirundo rustica* in Chernobyl, Ukraine. *Proceedings of the Royal Society, London*, 252, 51–57.
- Nelson, D. A. (1992). Song overproduction, song matching and selective attrition during development. In P. K. McGregor (Ed.), *Playback and studies of animal communication* (pp. 121–134). New York: Plenum Press.
- Owren, M. J., & Rendall, D. (1998). An affect-conditioning model of nonhuman primate vocal signaling. In D. Owings, M. Beecher & N. Thompson (Eds.), *Perspectives in ethology, Vol. 12: Communication* (pp. 299–346). New York: Plenum Press.
- 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.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology*, 26, 389–406.
- Owren, M. J., Seyfarth, R. M. & Cheney, D. L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and function. *Journal of the Acoustical Society of America*, 101, 2951–2963.
- Pallier, C., Christophe, A., & Mehler, J. (1997). Language-specific listening. *Trends in Cognitive Science*, 1, 129–132.
- Petersen, M. R., Beecher, M. D., Zoloth, S. R., Moody, D. B., & Stebbins, W. C. (1978). Neural lateralization of species-specific vocalizations by Japanese macaques. *Science*, 202, 324–326.
- Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *Journal of the Acoustical Society of America*, 66, 363–369.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow and Company, Inc.
- Premack, D. (1986). *Gavagai! or the future history of the animal language controversy*. Cambridge: MIT Press.
- Premack, D., & Premack, A. J. (1995). Origins of human social competence. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 205–218). Cambridge: MIT Press.
- Putnam, H. (1975). *Mind, language and reality*. Cambridge: Cambridge University Press.
- Quine, W. V. (1973). On the reasons for the indeterminacy of translation. *Journal of Philosophy*, 12, 178–183.
- Rendall, D., Rodman, P. S., & Edmond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, 51, 1007–1015.
- Robinson, J. G. (1979). An analysis of the organization of vocal communication in the titi monkey *Callicebus moloch*. *Zeitschrift für Tierpsychologie*, 49, 381–405.
- Robinson, J. G. (1984). Syntactic structures in the vocalizations of wedge-capped capuchin monkeys. *Cebus nigrivittatus*. *Behaviour*, 90, 46–79.
- Savage-Rumbaugh, E. S. (1986). *Ape language: From conditioned response to symbol*. New York: Columbia University Press.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, 58, 1–221.
- Schusterman, R. J., Gisiner, R., Grimm, B. K., & Hanggi, E. B. (1993). Behavior control by exclusion and attempts at establishing semanticity in marine mammals using match-to-sample paradigms. In H. L. Roitblat, L. M. Herman, and P. E. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 249–274). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Searcy, W. A. (1992). Measuring responses of female birds to male song. In P. K. McGregor (Ed.), *Playback and studies of animal communication* (pp. 175–190). New York: Plenum Press.
- Seyfarth, R. M., & Cheney, D. L. (1980). The ontogeny of vervet monkey alarm-calling behavior: A preliminary report. *Zeitschrift für Tierpsychologie*, 54, 37–56.
- Seyfarth, R. M., & Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, 34, 1640–1658.
- Seyfarth, R. M., & Cheney, D. L. (1997). Some general features of vocal development. in nonhuman primates. In C. T. Snowdon and M. Hausberger (Eds.), *Social influences on vocal development* (pp. 249–273). Cambridge: Cambridge University Press.
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210, 801–803.

- Shettleworth, S. (1998). *Cognition, evolution and behavior*. New York: Oxford University Press.
- Smith, W. J. (1977). *The behavior of communicating*. Cambridge: Harvard University Press.
- Snowdon, C. T., & Hausberger, M. (1997). *Social influences on vocal development*. Cambridge: Cambridge University Press.
- Soja, N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition*, 38, 179–211.
- Stebbins, W. C., & Sommers, M. S. (1992). Evolution, perception and the comparative method. In D. B. Webster, R. F. Fay, & A. N. Popper (Eds.), *The evolutionary biology of hearing* (pp. 211–228). New York: Springer-Verlag.
- Struhsaker, T. T. (1967). Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In S. A. Altmann (Eds.) *Social communication among primates* (pp. 281–324). Chicago: Chicago University Press.
- Studdert-Kennedy, M. (1998). The particulate origins of language generativity: From syllable to gesture. In J. Hurford, M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases* (pp. 202–221). Cambridge, UK: Cambridge University Press.
- Thompson, R. K. R. (1995). Natural and relational concepts in animals. In H. L. Roitblat & J. A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 175–224). Cambridge: MIT Press.
- Tinkelpaugh, O. L. (1928). An experimental study of representative factors in monkeys. *Journal of Comparative Psychology*, 8, 197–236.
- Todt, D., Hammerschmidt, K., Ansorge, V., & Fischer, J. (1995). The vocal behavior of barbary macaques (*Macaca sylvanus*): Call features and their performance in infants and adult. In E. Zimmerman, J.D. Newman & U. Jurgens (Eds.) *Current topics in primate vocal communication* (pp. 141–160). New York: Plenum Press.
- Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpenter, M. (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 35, 137–154.
- von Frisch, K. (1967). *The dance language and orientation of bees*. Cambridge: Belknap Press of Harvard University Press.
- Waser, P. M. (1977). Experimental playbacks show vocal mediation of avoidance on a forest monkey. *Nature*, 255, 56–58.
- Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature*, 382, 629–632.
- Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology*, 53, 205–214.
- Zahavi, A. (1993). The fallacy of conventional signalling. *Proceedings of the Royal Society, London*, 340, 227–230.
- Zahavi, A., & Zahavi, A. (1997). *The handicap principle: A missing piece of Darwin's puzzle*. Oxford: Oxford University Press.
- Zoloth, S. R., Petersen, M. R., Beecher, M. D., Green, S., Marler, P., Moody, D. B., & Stebbins, W. C. (1979). Species-specific perceptual processing of vocal sounds by monkeys. *Science*, 204, 870–872.
- Zuberbuhler, K., Noe, R., & Seyfarth, R. M. (1997). Diana monkey long-distance calls: Messages for conspecifics and predators. *Animal Behaviour*, 53, 589–604.