Reprinted from Primate Ontogeny, Cognition and Social Behaviour (eds. Else & Lee). Proceedings of the 10th Congress of the International Primatological Society vol. 3.

IV.5

Grading in the vocal repertoire of Silver Springs rhesus monkeys

E. H. PETERS

Introduction

The purpose of this chapter is to re-examine 'grading' as a characteristic of primate vocalizations and to suggest a new way of thinking about the organization of calls which may clarify the generation of extreme variety in vocal repertoire. The graded nature of primate calls was first reported by Rowell & Hinde (1962), following their pioneering application of sound spectrography to rhesus monkey calls. Working with sound samples from a captive troop of animals, these investigators defined a finite number of 'basic noises' but prefaced this list with the observation that there exists 'an almost infinite series of intermediates between the main sounds' and the suggestion that calls may 'intergrade with one another so that the variation is immense and it is impossible to give the full vocabulary written equivalents'. In a follow-up analysis of agonistic vocalizations, Rowell (1962) suggested that calls in this family could be arranged in a continuous graded series, with the particular points along the continuum identified and labeled in a somewhat arbitrary fashion.

Much of the difficulty which these and other investigators have encountered in the description of primate vocal repertoire may be related to overly strict adherence to the typological paradigm which emerges as an underlying feature of all such descriptive attempts. The standard procedure for describing primate sound signals continues to be that of defining a finite number of physically distinctive, relatively stereotyped acoustic units (i.e. the vocalization or call), each with its own meaning or message. This procedure can be taken to breathtakingly subtle levels of analysis, as Green's (1975) fine-grained description of Japanese macaque vocalizations has shown. The danger with an exclusively typological approach is that characteristics which cross-

link different calls may be obscured, making it more difficult to discern how pattern and meaning are actually generated and discerned by the animals themselves. The current practice of referring to the existence of 'intermediates' between the types or 'variants' as departures from the type may reify categories which have no such primacy of position in the mind of the animal.

Typological constraint in the description of primate sound signals may be related to the precedent set by the highly successful ornithologists, a group whose leadership in the analysis of animal sound is undoubtedly related to the salience of aural communication in birds. However, it is not always appreciated that the success of typological sound analysis for birds and many other animals is related to the prevalence of a high degree of stereotypy and contrast in their signals. These characteristics, as Morris (1957) has pointed out, enhance signal clarity and decrease ambiguity. Given the perceptual-cognitive limitations shown to characterize many species of animals, such stereotypy and high contrast might be necessary for reliable signal processing. Suppose, however, evolutionary processes produced an animal with the perceptual acuity to process incoming sound signals even when the important cues are subtle in nature and embedded in a miscellany of extraneous stimuli. Suppose this animal had the cognitive complexity to synthesize simultaneously the information in multiple cues. Suppose it was sensitive to small changes in meaning wrought by small changes in signal form. In short, given an appropriate level of perceptual-cognitive complexity, neither stereotypy nor high contrast may be necessary for reliable processing of conspecific sound signals. For any species, the nature of its sound signal system is likely to be correlated with its perceptual-cognitive capacity. As this capacity shows evidence of a change in evolutionary grade, it is reasonable to look for emergent properties in the co-evolving signal system. It is also reasonable to suggest that the same system of signal description which works well with birds may not work quite so well with monkeys.

Methods

The data on which my own analysis of rhesus monkey vocalizations is based were collected from a troop of feral rhesus monkeys which inhabit the lowland floodplain adjacent to the Silver River near Silver Springs, Florida. This troop is part of a colony of several hundred animals descended from monkeys released into the wild in the 1940s as an exotic and seemingly natural wildlife exhibit. Record-

ings of vocalizations and associated behavior were collected between October 1978 and May 1979 from a single troop of approximately 30 individually identifiable animals. Spectrograms of vocalizations were made on a Kay Sonagraph model 6061A using a 150 Hz filter to scan the 80–80 000 Hz range.

Results

It was while attempting to sort vocalization samples from the above monkeys using the standard typological paradigm that difficulties with this procedure became apparent. These included problems both in defining what Altmann (1967) has called 'criteria of sequential demarcation' (given a temporal continuum of sound, where does a specific call begin and where does it end?) and in defining 'criteria of class membership' (when are two sounds sufficiently similar so that they should be considered members of the same class?).

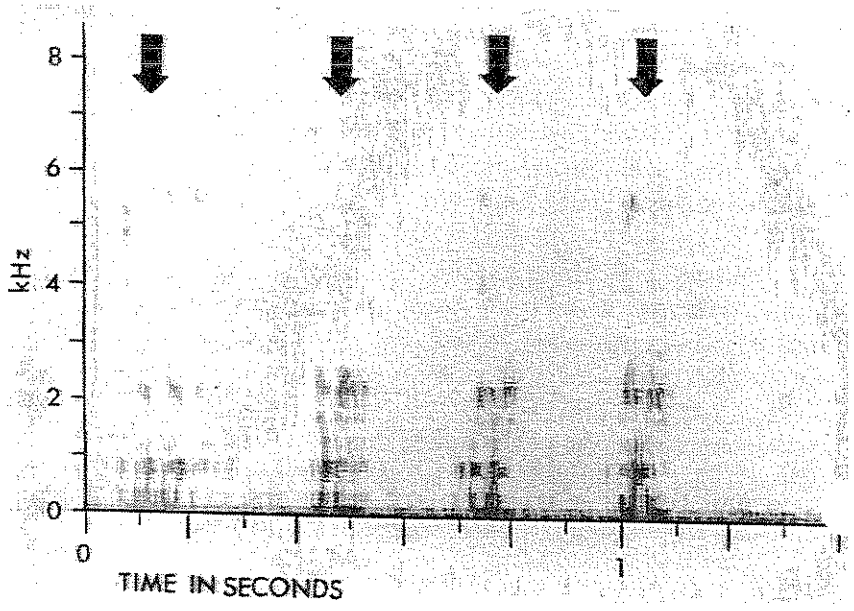
For many animal sound signals the beginning and the end of a call can be easily and reliably determined. When a single unit of sound (unit, as defined by Struhsaker (1967), refers to a continuous tracing along the temporal axis of the sonogram) is surrounded by substantial boundaries of silence, then that single unit can be reliably understood to be the call. Rhesus tonal coos and many barks and grunts (noisy, one-unit sounds with either a high pitch or a low pitch, respectively) fall into this category. Other species-typical sound signals are made up of multiple units, but if the component units are similar, the time interval between units fairly regular, and the time interval between component units shorter than the time interval between groupings of units, then it is again relatively easy to distinguish the beginning and the end of the call. The geckering (Rowell & Hinde, 1962) of immature rhesus and the chortling (Lindburg, 1971) of adult and juvenile females are examples of multi-unit calls for which sequential demarcation between individual calls is still fairly obvious.

Problems in defining criteria of sequential demarcation exist for calls composed of multiple units with irregularly shaped, highly variable components and irregular spacing between units. The family of calls which Rowell (1962) called 'agonistic' often includes examples which fall into this category. It includes high-pitched screams (Gouzoules, Gouzoules & Marler (1984) have called these 'aid enlistment' calls) and low-pitched threat calls. Many of these fear- and threat-expressing calls are composed of multiple units, with component units often of dissimilar form, number and/or spacing. The response to a single stimulus situation might be a train of irregular

sound units which are several seconds long. How much of this train should be considered a single call? Given the highly variable nature of the component units and the resultant whole, how can any two sampled calls be considered members of the same class or type? The difficulty and the importance of this last issue has been underscored by the recent discovery of Cheney & Seyfarth (1982) that even spectrographically similar-looking vervet monkey grunts can function in distinctly different ways.

Recently Gouzoules *et al.* (1984) have made some progress in this issue by defining repeatable features found in 'aid enlistment' screams. For example, screams which contain a wide bandwidth of noise ('noisy screams') were classified separately from those with energy restricted to a narrow range ('tonal screams'), regardless of the number of units, the shape of the units or the spacing between them. As useful as this approach is as an essential first step, it is only the beginning of an explanation. Fully 25% of the uninterrupted vocal units which Gouzoules *et al.* sampled either could not be classified or consisted of characteristics representative of more than one of their defined scream classes. Their analysis also gives no indication of the

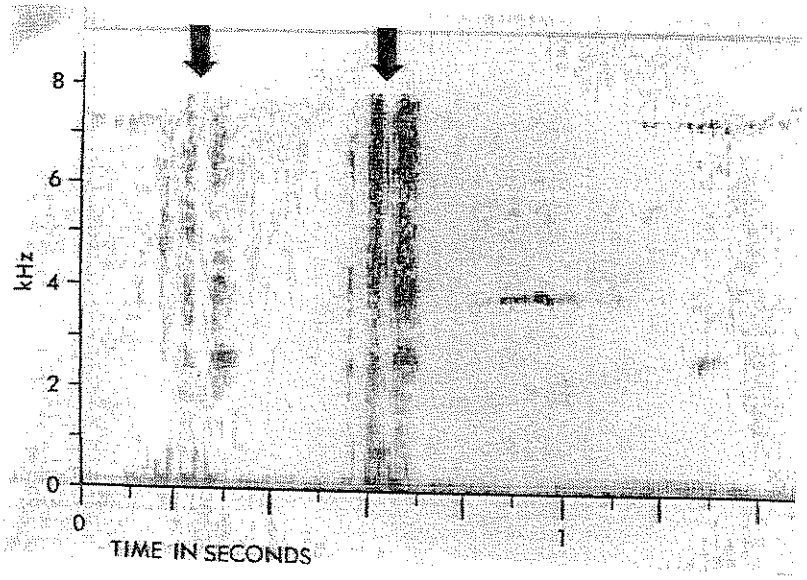
Fig. 1. A multi-unit threat call composed of four distinctive units each showing evidence of vertical striation. Arrows clarify the cadence discernible to the human ear. Note that energy emphasis is below 3000 Hz in this particular call.



fact that units of different classes may be combined into long sequences of response to a single stimulus situation. All of this seems to indicate that rhesus vocal skill encompasses something more than a simple typological repertoire.

My own consciousness about rhesus vocal behavior was raised when I began to think about a missing element in Rowell & Hinde's (1962) cataloguing of rhesus threat calls. Specifically, these investigators distinguished between a one-unit threat call with vertical striation (the 'growl') and a one-unit threat call without vertical striation (the 'bark'). These in turn were categorically separated from all multi-unit threat calls which they called 'pant threats'. Now multi-unit threat calls can vary enormously in the number of units and also vary somewhat in the size, shape and spacing of the units, in pitch and in volume. They can also show vertical striation as assuredly as the one-unit 'growl' (see Fig. 1). Consistency in the use of cataloguing criteria would therefore demand that striated multi-unit threat calls be categorically distinguished from unstriated multi-unit calls just as Rowell & Hinde (1962) did for one-unit threat calls. On the other hand, if one-unit threat calls were separable from multi-unit calls, why not distinguish two-unit calls from three-unit calls, etc? Each of these in turn would have to be split into striated (Fig. 2) and unstriated

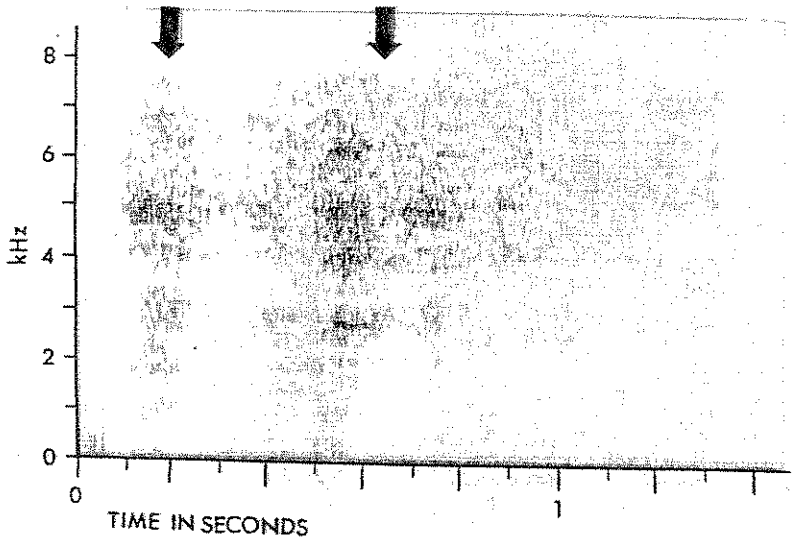
Fig. 2. A two-unit threat call with well-defined vertical striation (ignore background noise after 800 ms). Note that energy is broadly distributed to 8000 Hz and beyond.



versions (Fig. 3). The potential for splitting generated by the logical extension of a few simple cataloguing criteria makes it desirable to seek a more parsimonious way of conceptualizing this immense variety.

In threat situations, the two parameters under examination, number of units and vertical striation, seemed to be varied independently by the monkeys. Furthermore, each parameter showed evidence of existing as a continuum. Number of units could typically range from one to seven (and sometimes more). Vertical striation could be well defined (Fig. 2), partially visible (Fig. 1) or completely absent (Fig. 3). Volume, cadence, and the length of and distribution of energy in the constituent units also showed evidence of being both variable along a continuum and at least somewhat independently manipulable (note, for example, differences in the frequency distribution of energy in Figs 1, 2 and 3). Given the observation that the form of any specific threat call could vary enormously depending on the specific, separable contribution of each of the component-form gradients, then a new conceptualization of the physical basis of call variety becomes possible. Stereotyped calls can be conceived of as calls for which selection has produced a typical and relatively invariant position along each of the several contributing form gradients (e.g. number of units, degree of striation, volume, pitch, cadence, length of constituent units, etc.). On the other hand,

Fig. 3. A multi-unit threat call which the human ear perceives as composed of one short and one longer unit. This two-unit call lacks evidence of vertical striation. Note energy emphasis around 5000 Hz.



'graded' calls and extreme variety of vocal repertoire result when *more than one* position along each of the contributing form gradients is possible.

Discussion

The existence of component-form gradients and the possibility that position along each gradient makes a separate contribution to the emergent form of any given call leads one to ask whether semantic gradients accompany the form gradients (i.e. whether component-meaning gradients, not just categorical shifts in meaning, characterize rhesus monkey calls). Conceivably, the meaning of a specific vocalization may be the result of the convergence of effects wrought by position along all component form-meaning gradients. Thus each call might be customized both in form and meaning and any attempt to set up typological call categories would only weakly display the richness of the system. In practice, some typologizing has proven to be a useful way to understand information transfer in non-human primates. Just as assuredly, however, the immense variability of the individual sound signals produced by certain species of non-human primates has continued to thwart comprehensive typologizing and to suggest a more flexible, multi-faceted system of signal generation.

The typologizing habit seems to be an ancient thread in Western civilization (witness Plato's concept of ideal forms) and it may even be a standard cognitive tool of the human mind. Despite its proven usefulness, however, it is not necessarily the only way to conceptualize the organization of animal sound signals and it may not always be the best way to understand what is going on. A better understanding of both graded and non-graded signals (and of the evolutionary relationship between them) may emerge if we loosen this strictly typological approach and try to include some understanding of component gradients and their potential for generating enormous variety in signal structure and meaning. Perhaps the future study of non-human primate vocalizations will not just be a search for the meaning of a finite number of categorical call types but will also be a search for contributing gradient constituents of large call families and for any change in meaning wrought by change in position along each contributing gradient.

References

Altmann, S. A. (1967) The structure of primate social communication. In *Social Communication among Primates*, ed. S. Altmann, pp. 325–62. Chicago: University of Chicago Press

Cheney, D. L. & Seyfarth, R. M. (1982) How vervet monkeys perceive their grunts: field playback experiments. *Anim. Behav.*, **30**(3), 739–51

Gouzoules, S., Gouzoules, H. & Marler, P. (1984) Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Anim. Behav.*, **32**, 182–93

Green, S. (1975) Communication by a graded system in Japanese monkeys. In *Primate Behavior*, Vol. 4, ed. L. A. Rosenblum, pp. 1–101. New York: Academic Press

Lindburg, D. G. (1971) The rhesus monkey in north India: an ecological and behavioral study. In *Primate Behavior: Developments in Field and Laboratory Research*, ed. L. A. Rosenblum, pp. 1–106. New York: Academic Press

Morris, D. (1957) 'Typical intensity' and its relation to the problem of ritualization. *Behaviour*, **11**, 1–12

Rowell, T. E. (1962) Agonistic noises of the rhesus monkey (*Macaca mulatta*). *Symp. Zool. Soc. London*, **8**, 91–6

Rowell, T. E. & Hinde, R. A. (1962) Vocal communication by the rhesus monkey (*Macaca mulatta*). Proc. Zool. Soc. London, **138**, 279–94

Struhsaker, T. T. (1967) Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In *Social Communication among Primates*, ed S. A. Altmann, pp. 281–324. Chicago: University of Chicago Press