Eur Arch Psychiatr Neurol Sci (1986) 236:35–39

European Archives of Psychiatry and Neurological Sciences

Interrelations Between Structure and Function in the Vocal Repertoire of *Saimiri*

Asking the Monkeys Themselves Where to Split and Where to Lump

M. Maurus, K.-M. Streit, D. Barclay, E. Wiesner, and B. Kuehlmorgen

Max-Planck-Institut für Psychiatrie, Kraepelinstrasse 2, D-8000 München 40, Federal Republic of Germany

Summary. The identification of units used in primate vocal behaviour and the isolation of components essential for intraspecific communication (i.e. having a function) from possibly inessential ones (i.e. having no function) is an absolute prerequisite to a successful analysis of primate vocalizations. For most calls special methods are required because neither the essentials of structure nor their relation to function are known, nor the context involved. For calls without any distinct frequency modulation we showed that the animals utter calls containing categorical amplitude changes. The changes within one call are either very small or very large. The positions or the very large amplitude changes within a call depend on the social situation the animals were subjected to and affect the immediately succeeding vocalization of the conspecific. This proves that in vocal communication discrete components are produced (structure) as well as used (function).

Key words: Primate vocalization – Discrete units – Amplitude changes – "Dialogues"

When analyzing the acoustic communication of non-human primates, it has become common practice that observers select "call types" primarily on the basis of the auditory impression of the sound and on visual characteristics as derived from sonagrams. This procedure can result in conflicting catalogs for one and the same species (e.g. for the rhesus monkey, Reynolds 1976; for the squirrel monkey, Newman 1985). Logically, there can only be one correct catalog, but it is impossible to clarify whether it is one of the catalogs presented. In recent studies, it has been demonstrated, at least for some "call types", that observer decision was not in accord with the monkey's use of acoustic signals (e.g. Cheney and Seyfarth 1982; Masataka 1983). These findings consequently would make it necessary to either re-examine all "call types" based on observer decision by additional experimental and equipment means (e.g. playback, call manipulation, call analysis) or apply procedures not based on observer decision but on "asking" the monkeys themselves.

The re-examination of "call types" by playback and call manipulation experiments has proved to be applicable and successful only where warning, contact and mother-child behaviour was concerned. Excluding observer decision, on the other hand, reduces the risk of bias originating from previously established behavioural units. It is applicable also when structure or function are unknown.

This is a crucial point for, in a wide area of intra-specific communication, there is primarily no other observable object for analysis than a sequel of vocal utterances with unknown functions. Only the structure of the call can be resolved into its component parts in a first step of analysis.

Vocal utterances can be described by using the parameters frequency, amplitude and time. In most cases, however, it cannot be perceived intuitively by the observer which of these parameters, or interrelations of them, are connected with components essential for transmitting information to the conspecific. Exceptions are contact and warning calls, for their function is made evident by the context in which they occur. In other cases, the use of statistical methods may be helpful. There could be, for instance, structural components which are produced more often than others. Preliminarily, they may be assumed to be representative of communicative functions. This assumption seems justified by the following paradigm.

In human speech, in plosive sounds like b and p, for instance, the distinctive component is the voice onset time (e.g. Liberman et al. 1967). This means that, on one side of the categorical boundary, the sound is perceived as b, on the other side as p. Voice onset time is also produced categorically thus reducing ambiguity in perception. "Normal" adults, when producing b and p, have their voice onset time values statistically distributed over an area sufficiently distant from the categorical boundary (e.g. Sweeting and Baken 1982). Supposing that non-human primates also use categorical components in communication, one might expect accumulations in the statistical distribution of the physical parameters or combinations thereof. A search for such accumulations in the acoustic repertoire of the monkeys would therefore seem promising. How this first step in the analysis may be taken, will be explained below.

1. Analysis of Structure

In order to avoid bias (e.g. over-emphasis on frequency-time domain as suggested by relying on sonagrams), the same degree of importance should be attached to all parameters and their combinations (e.g. course of frequency, course of ampli-

---

Offprint requests to: M. Maurus at the above address

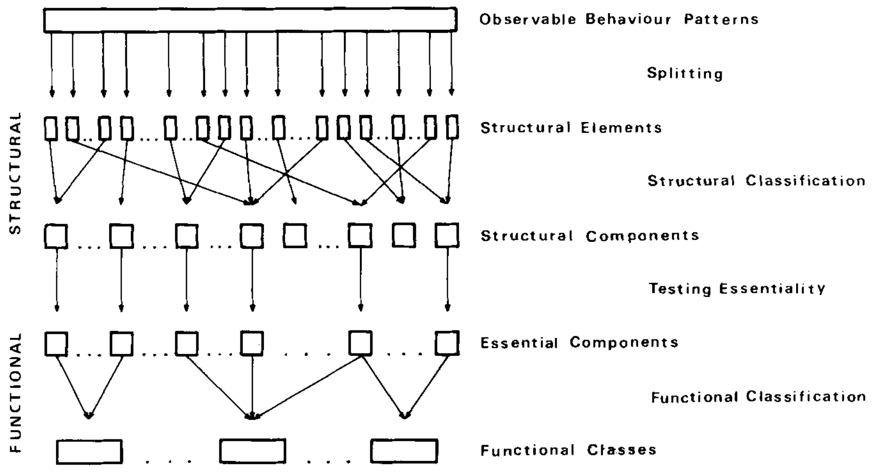

Fig. 1. Schematic representation of procedures for analysing intra-specific communication

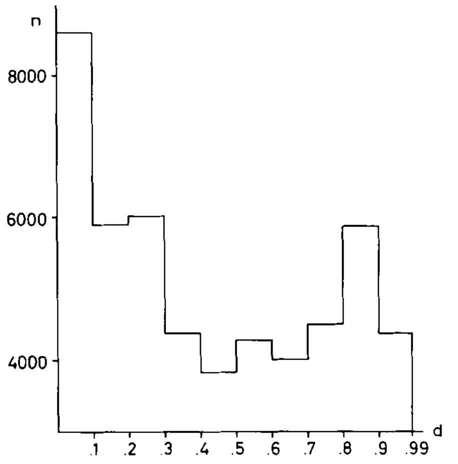

Fig. 2. Histogram of relative differences, d, between the amplitude maxima of immediately succeeding segments (values were normalized to the greater of the two maxima). Data from x calls uttered by the 5 animals T, P_1–P_4. The upward difference to the first segment of the call and the downward difference from the last segment of the call were not included (hence excluding differences with value 1.0) (P_5 was not considered as she did not utter any x calls)

tude, power spectrum, course of power spectrum). This may make it necessary, when performing the analytical steps shown in Fig. 1, to apply different analysis procedures to different parameters. In the following, only one parameter combination has been chosen for explaining the procedure, namely, the course of the amplitude in squirrel monkey calls not containing elements distinctly modulated in frequency (this category is denoted by x; the category of calls containing elements distinctly modulated in frequency is denoted by z). For splitting (Fig. 1), the calls were segmented in accordance with the course of the amplitude envelope. Each minimum in this course defines a boundary between two segments. Each of the segments produced this way had one maximum amplitude between start and end. For each segment, time, amplitude and frequency parameters were calculated and used in describing the segments which subsequently could be used for describing the entire call consisting of a sequel of these segments (frequency parameters are not considered in this paper).

In order to find accumulations in the statistical distributions of the calculated parameters, a classification procedure was applied that groups similar segments according to the values of various multidimensional descriptions of the segments (for further details see Maurus et al. 1984). It was found that there are segment classes with a relatively high amplitude maximum (in relation to the absolute maximum in the call), segment classes with a relatively low amplitude maximum, but no segment classes with a medium amplitude maximum. Consequently, the amplitudes of successive segments in a call are most often either very similar or very different (Fig. 2). In this way no, one or more amplitude leaps (i.e. large amplitude differences between succeeding segments) may be located at different position within the call.

The categorical production of minor amplitude changes and amplitude leaps suggests that this aspect of the amplitude modulation may refer to a component essential for intra-specific communication. If so, it would bring up the question what makes the amplitude leaps essential to communication. The mere fact of their occurrence in a call? The number of leaps per call? The position of the leaps within the call? Or a more complex combination of number and position? In order to find an answer to at least one of these questions, the next step in the analytical process is required.

2. Analysis of Function

To prove that a structural component has a function in intraspecific communication, it must be demonstrated that a conspecific "responds adequately" to this component, or that the component is an "adequate response" to a conspecific's preceding signal. We found that animals produce calls containing different numbers of amplitude leaps when vocalizing in the presence of different conspecifics (Fig. 3). The example in Fig. 3 refers to all calls of type x that contain at least one section beginning with two or more immediately succeeding high amplitude segments. With regard only to that part following the first occurrence of such a section within the call, these calls were arranged in three classes. The first class, x_1, comprised all calls ending with that section. The calls of the second class, x_2, contained at least one low amplitude segment succeeding the section. Calls of the third class, x_3, differed from those of class two by being additionally followed by at least one high amplitude segment.

Figure 3 illustrates the frequencies of occurrence of the calls of classes x_1, x_2 and x_3 uttered by the test animal T in the

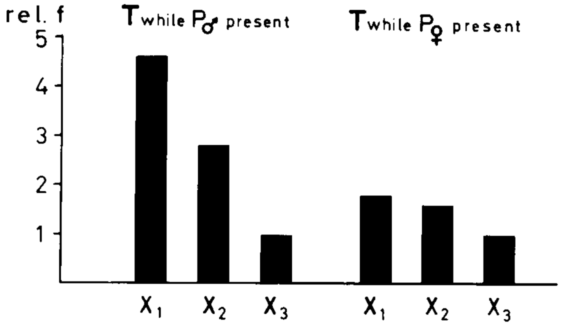

Fig. 3. Frequency of occurrence of x_1, x_2, x_3 uttered by test animal T in the presence of male and female conspecific, respectively. In both cases the frequency of occurrence of x_3 was set to 1

presence of either a male or female conspecific. This result shows that the production of amplitude leaps at particular positions within the call is influenced by the social situation (for details see Maurus et al. 1984). It may also indicate, but does not prove, that this component has a communicative function. To prove it, it must be shown that this component affects the conspecific's behaviour. But the point is how to prove it, for it cannot be stated a priori that a call and the conspecific's succeeding vocalization are functionally interrelated. If monkey vocalizations are an effect of the animal's internal state exclusively, as stated by Lancaster (1975), for instance, the utterances of one animal cannot be claimed to be related to those of a conspecific vocalizing simultaneously. Both animals may experience different internal states independently from each other. Only for warning, contact and mother-child behaviour is there evidence that when squirrel monkeys communicate acoustically, the conspecific "responds adequately" to an immediately preceding call. But before taking analytical steps based on partner "reaction", it must be proved for the entire repertoire that one animal's call and the conspecific's succeeding vocalization are functionally interrelated.

For this purpose, a test animal T was placed in a cage where it was joined by one of five partners (P_1–P_5) consecutively. The calls of both animals were tape-recorded telemetrically (Maurus and Szabolcs 1971). All vocalizations uttered by the animals were arranged in the two categories x and z according to unambiguous structural criteria only (for more details see Maurus et al. 1985). This procedure possibly has the effect that a number of unknown functions is put into each of the structural categories. A structure-based categorization not in accordance with function may blur the results but not markedly decrease the reliability of results provided reproducibility is assured.

Evaluation of intra-individual sequels (frequency of transitions from the preceding to the immediately succeeding call: xx; xz; zx; zz) revealed that T's vocalization was less affected when the partner was silent than when it vocalized (Fig. 4). The fact that a silent partner does not exercise as much influence as a vocalizing one suggests a dependence of vocalizations by T on those of P. But what causes the dependence? Is it the fact that the partner vocalizes or what he vocalizes? We checked whether T vocalized differently after different calls from P. For this purpose, we reviewed all instances in which one call was uttered by P between two calls from T (T-P-T sequels; calls of partners parenthesized: (x); (z)), when, e.g. an x call from T was followed by an (x) from P which T responded to with an x: x(x)x etc. In response to an (x) call from P following an x from T, T may utter an x or a z call: x(x)x or x(x)z. This would result in a specific ratio of x(x)x to x(x)z for all sequels beginning with x(x): $\frac{x(x)x}{x(x)z}$.

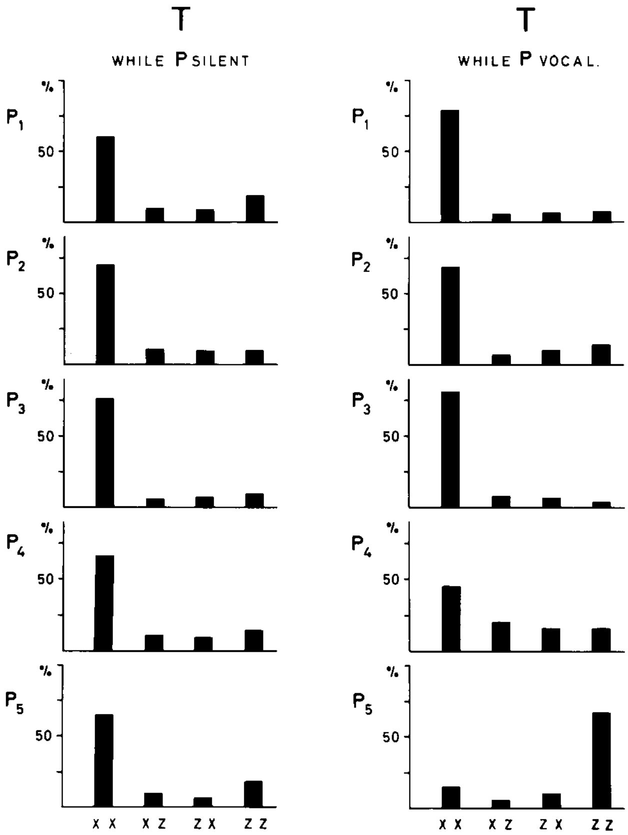

Fig. 4. Relative frequencies of calls uttered by T. *Left:* calls uttered by T while P_1–P_5 were silent. *Right:* calls uttered by T while P_1–P_5 vocalized. ("... while ... silent [vocal.] means that between the starts of two consecutive calls of one animal, the other animal uttered no call [at least one call])

A change in the ratio due to P uttering a (z) instead of an (x) call, i.e. $\frac{x(x)x}{x(x)z} \neq \frac{x(z)x}{x(z)z}$, is an indication that the vocalization from P influences the following vocalization by T, and the greater the deviation, the stronger the influence.

Shown in Fig. 5 (left) are the ratios for all instances in which an x call from T was succeeded by an (x) or (z) call from P, or a z was followed by an (x) or (z). Listed on the right are all instances in which an (x) call from P was preceded by an x or z call from T, or a (z) was preceded by an x or z. In most instances the ratios differed considerably but changed with each partner. According to these data, structural categorization produced by the partners has an effect on the test animal's vocalizations.

The results depicted in Fig. 4 suggest differently constructed call sequences for each animal depending on whether the conspecific vocalized or was silent, and it seems that mainly vocalizing conspecifics influence the vocal behaviour of their partners – an indication that many vocal changes are attributable to acoustic, and not (or not only) to non-acoustic signals.

The influence exerted by partner vocalizations is demonstrated even more clearly by the results shown in Fig. 5: the

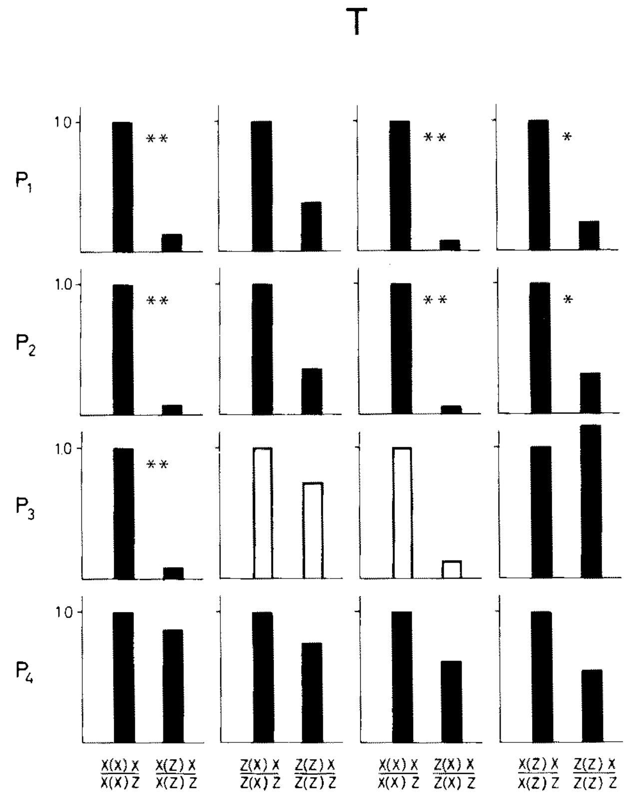

Fig. 5. Ratios of x to z in T "replies" $\left(\frac{x}{z}\right)$ after an (x) $\left(\frac{(x)x}{(x)z}\right)$ or a (z) $\left(\frac{(z)x}{(z)z}\right)$ from P_1–P_4 when the call from P_1–P_4 was preceded by an x from T: $\frac{x(x)x}{x(x)z}$ (1st column from left); $\frac{x(z)x}{x(z)z}$ (2nd column from left; and when preceded by z from T: $\frac{z(x)x}{z(x)z}$ (3rd column from left); $\frac{z(z)x}{z(z)z}$ (4th column from left). Columns 5–8 from left: different pairwise comparisons of these ratios. Significance tests (χ^2-test, $df = 1$): * ($P < 0.05$), ** ($P < 0.005$). White columns: Results of such a small number of events that a test did not seem reasonable

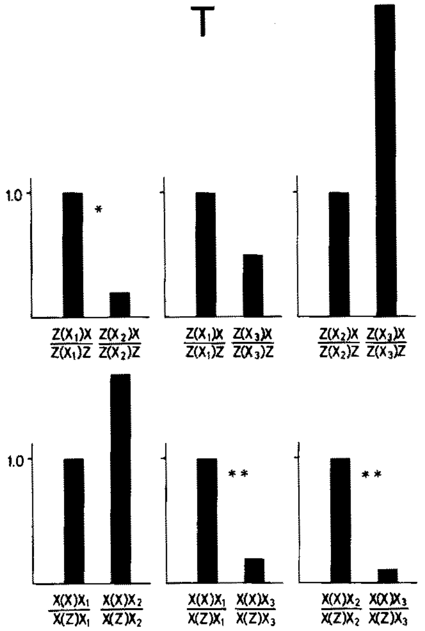

Fig. 6. *Upper part:* Ratios of x to z in T "replies" $\left(\frac{x}{z}\right)$ after an (x_i) $\left(\frac{(x_i)x}{(x_i)z}\right)$, $i = 1, 2, 3$ from P when the call from P was preceded by a z from T: $\frac{z(x_i)x}{z(x_i)z}$. *Lower part:* Ratios of (x) to (z) from P preceding an x_i from T $\left(\frac{(x)x_i}{(z)x_i}\right)$, $i = 1, 2, 3$ when the call from P was preceded by an x from T: $\frac{x(x)x_i}{x(z)x_i}$. Significance tests: χ^2-test, $df = 1$, for the pairwise comparisons x_1, x_2; x_1, x_3; x_2, x_3: * ($P < 0.05$), ** ($P < 0.005$). P means each of P_1–P_4; all experiments were summarized

test animal's reaction depends not only on the category of the partner call immediately preceding its "answer" but also on the call it uttered immediately before the partner's call. Although the differences noted were not the same for all animals tested, the results of our experiments revealed that in communicating acoustically, squirrel monkeys utter calls which influence the vocal behaviour of conspecifics, and that the degree of influence depends on the structural category of the call. In other words, the "answer" is affected by the call uttered immediately before.

Based on these findings, we can now question whether the amplitude leap is a component essential for intra-specific communication. This can be examined by taking into consideration the arguments related to Fig. 5. Since x and z are functionally different, the relation between their frequencies of occurrence can serve to detect differences in function of further structural categories. This x/z relation can be used in T-P-T sequels, considering calls from T (Fig. 6, upper part) as well as calls from P (Fig. 6, lower part). When, in T-P-T sequels, T began with a z followed by either (x_1), (x_2) or (x_3) from P, then T's x/z ratios differed significantly with respect to (x_1) and (x_2) only (Fig. 6, upper part). When, in T-P-T sequels, T began with an x followed by either an (x) or (z) from P and T "responded" with either x_1, x_2 or x_3, an x_1 followed an (x) or (z) with significantly different ratio than x_3 (resp. x_2 than x_3). This means that x_3 differs from x_1 and x_2 when it is the last call of a T-P-T sequel succeeding an (x) or (z), while (x_1) and (x_2) are responded to differently by T when uttered by P. The three categories affect the conspecifics' vocalizations differently. This effect apparently depends on the circumstances in a very complex way. The communicative function seems to be influenced by the call's position in the sequel. In human language, such influence is part of grammar. The role it plays in primate communication remains to be investigated.

Even though we do not yet know every detail, we have demonstrated, by using procedures independent of observer decision, that amplitude leaps are essential for intra-specific communication. Other structural components can be investigated in a similar manner so that finally the catalog of components used by non-human primates in acoustic communication may be completed. This catalog may contribute to our understanding of primate communicative processes and also increase our knowledge of the phylogeny of human speech. Our approach is not an attempt to teach apes human sign language in order to explore their cognitive abilities (e.g. Terrace 1984). It is a search for basic elements in human speech which exist, in final or preliminary form in the acoustic communication of non-human primates.

References

Cheney DL, Seyfarth RM (1982) How vervet monkeys perceive their grunts: field playback experiments. Anim Behav 30:739–751

Lancaster J (1975) Primate behavior and the emergence of human culture. Holt, Rinhart and Winston, New York

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy MG (1967) Perception of speech code. Psychol Rev 74:431–461

Masataka N (1983) Categorical responses to natural and synthesized alarm calls in Goeldi's monkeys *(Callimico goeldii)*. Primates 24:40–51

Maurus M, Szabolcs J (1971) Kleinstsender für die Übertragung von Affenlauten. Naturwissenschaften 5:273–274

Maurus M, Streit K-M, Geissler B, Barclay D, Wiesner E, Kuehlmorgen B (1984) Categorical differentiation in amplitude changes of squirrel monkey calls. Lang Commun 4:195–208

Maurus M, Kuehlmorgen B, Wiesner E, Barclay D, Streit K-M (1985) "Dialogues" between squirrel monkeys. Lang Commun 5: 185–191

Newman JD (1985) Squirrel monkey communication. In: Rosenblum LA, Coe CL (eds) Handbook of squirrel monkey research. Plenum Press, New York, pp 99–126

Reynolds V (1976) The origins of a behavioural vocabulary: the case of the rhesus monkey. J Theory Soc Behav 6:105–142

Sweeting PM, Baken RJ (1982) Voice onset time in a normal-aged population. J Speech Hear Res 25:129–134

Terrace HS (1984) "Language" in apes. In: Harré R, Reynolds V (eds) The meaning of primate signals. Cambridge University Press, Cambridge, pp 179–207

Received December 10, 1985