

COMPREHENSION OF VOCALIZATIONS ACROSS SPECIES

Reijo Aulanko*, Lea Lelnonen, Ilkka Linnankoski and Maija Laakso

*Department of Phonetics, University of Helsinki, Finland
Department of Physiology, University of Helsinki, Finland

ABSTRACT

This is a study of how well naive human listeners can interpret the vocalizations of stump-tail macaques (*Macaca arctoides*). The monkey vocalizations were recorded in different behavioural situations. The human listeners were asked to classify 18 vocalizations into one of 7 semantic categories. The listeners were quite unanimous in their judgements, which indicates that they may have based their interpretations on some kind of common "feature analysis" of the vocalizations. The interpretations of the listeners were also mostly "correct", i.e., the listeners were able to infer the situation in which the monkey had produced the sound. These results may be taken to suggest a possible common basis for the vocal behaviour of all primates.

1. INTRODUCTION

It is well known from everyday life that people and their pet animals can understand each other (or that they at least seem to reach a consensus on certain issues). The owner of a cat or a dog should find it easy to make very accurate interpretations (according to personal judgement) of the behaviour of the pet, e.g. of its vocalizations. Similarly, an animal sometimes reacts to the speech of its human companion as if it understands the human language. These cases are not, however, indications of *language* comprehension in the strict sense. The reactions of an animal are determined primarily by all kinds of non-verbal cues, and the most important phonetic aspects of the human speech are its *prosodic* characteristics — rather than the purely phonological structure of the utterance.

To put it simply, comprehension of a vocal message is an interpretation or understanding of the "internal state" of the sender. The correctness of the interpretation can be inferred from the reaction of the receiver. Humans can react verbally, but in the case of other species we have to deduce the interpretation of the message only on the basis of other kinds of overt (non-verbal) behaviour.

Because of their common evolutionary history, the basic mechanisms of sound production are similar in all mammals. There are similarities in the vocal apparatus as well as in the neural control of behaviour. Vocalizations of non-human primates are taken to be mainly reflections of their emotional-motivational state. In human speech, indications of such 'internal' states are often conveyed by prosodic or paralinguistic features. There may be enough acoustic similarity in the emotional-motivational vocalizations of human and non-human primates for comprehension across species.

In human speech, the various emotional and motivational states are reflected primarily in the general voice quality and the prosodic characteristics of speech, i.e. pitch, rhythm, and loudness (e.g. [4, 7]). These auditory characteristics normally co-occur with different kinds of facial expressions and body movements, but the auditory cues are usually sufficient for the identification of the speaker's emotional state.

Human beings are used to inferring the emotional state of a speaker from the acoustic characteristics of his/her speech. An interesting question would be how well these "inference rules" can be applied to the vocalizations of another species.

2. AIM

In this study [3], we explored the ability of the representatives of one species to interpret the vocalizations of another species. More specifically, we tried to determine how well naive human listeners can interpret the vocalizations of another primate species, viz. stump-tail macaques (*Macaca arctoides*). "Interpreting" is here defined as identifying the emotional-motivational state of the monkey during the production of different sounds.

The ultimate aim in studies like this is to resolve the question of a possible common control of emotional-motivational vocal behaviour in mammals. In other words, we are looking for universals in communicative behaviour.

3. RESEARCH MATERIAL

Sounds. Recordings of the macaque vocalizations were made in many different behavioural situations at the Department of Physiology, University of Helsinki, in the colony of stump-tail macaques (*Macaca arctoides*) at present consisting of 12 monkeys (Marantz CP430 tape recorder, AKG C 568 EB microphone). On the basis of the situation and the total behaviour of the monkey, the sounds used in this study were taken to represent seven different categories of psychological states: (1) aggression, (2) fear, (3) sexual arousal, (4) dominance, (5) submission, (6) contentment, (7) calling / informing (contacting). The criteria used in this classification were based on, e.g., the posture and facial expressions of the monkeys, as they are generally used in primate behavioural studies [1, 2].

The vocalization sequences were digitized and tapes for the listening test were prepared, where the vocalizations occurred in a random order. The vocalizations in the test material were analyzed acoustically using sound spectrograms and computerized FFT spectra (Fig. 1). The acoustic characteristics of the vocalizations are described elsewhere [3].

Listening test. Eighteen sound sequences ("whole vocalizations") were selected from all the recorded material for the listening test. The 18 sounds represented different behavioural situations.

A total of seventy-five subjects (50 women and 25 men) participated in the listening test. They were 19–62 years of age, most of them students (of medicine,

dentistry, and anthropology), but there were also some speech therapists, medical doctors, technicians, and nurses. The subjects were not familiar with the vocalizations of the *Macaca arctoides*, but 43 of the 75 listeners had daily contacts with domestic or pet animals.

In a forced-choice test, the subjects were asked to classify each vocalization into one of the seven response categories, each of which was described by (the Finnish equivalents of) the following adjectives:

1. angry, cross, raging
2. frightened, timid, terrified
3. ecstatic, excited, orgasmic
4. commanding, threatening, domineering
5. submissive, pleading, begging
6. satisfied, satiated, delighted
7. calling, informing, addressing

The subjects were given two minutes to become acquainted with the response classes by thinking about each adjective momentarily.

The sounds were presented in a random order. Five of the 18 sounds were included twice in the test tape in order to find out the replicability of the subjects' classifications. (Thus, there was a total of 23 sounds to be judged.) The subjects heard a sound sample twice before a 10-second response interval during which they had to write down the number of the response class that best characterized the sound.

4. RESULTS

4.1. Listener agreement

The responses were not distributed randomly, i.e., the subjects were quite unanimous about the "meaning" of most of the monkey vocalizations (Table 1). The most variable responses were elicited by the vocalizations produced by aggressive monkeys, whereas the listeners were most unanimous in their responses to the "dominance roar" of the leading male.

4.2. "Correct" interpretations

On the average, 60 per cent of the listeners' interpretations were "correct". A response was defined as correct when it corresponded to the original behavioural classification of the monkeys' vocalizations. Somehow the listeners could infer the situation where the sound had been produced.

It has to be noted that certain sounds were originally taken to reflect at least two behavioural classes simultaneously. For example, one of the vocalizations was produced by a female macaque in a situation where her non-vocal behaviour indicated both aggression and fear. These alternatives have not been taken into account in Table 1, where only the responses falling on the diagonal are considered "correct", on the basis of the primary characteristic of the monkey's behaviour.

There were no general differences between the interpretations of men and women, or between those of younger and older listeners, although some individual vocalizations were classified differently. In contrast, daily contact with animals had a significant effect: those listeners who owned animals had more correct answers than those who did not have pets at home (61.5 % vs. 56.8 %).

The sounds that were best identified were a vocalization of a female monkey associated with pleasure (85 % of the subjects had the correct classification) and a dominance roar of a dominating male monkey (84 %). All meaning categories included vocalizations that were classified correctly by more than half of the listeners. Most of the subjects gave the right answer to 13 different vocalizations. One vocalization (threat grunts of a female monkey) was misclassified by all subjects. The distributions of the classifications of the five vocalizations that were presented twice remained stable.

5. DISCUSSION

The rather high general agreement among listeners shows that human listeners do tend to interpret monkey vocalizations that they have never heard before on the basis of some common ideas about the effects of different emotions on the sound production of another primate species.

The high proportion of "correct" interpretations shows that there are acoustic characteristics in the vocalizations which enable naive human listeners to infer the emotional-motivational state of the vocalizing macaque.

The most plausible explanation of the ability to interpret these monkey vocalizations is that the effects of different emotional and motivational states produce rather similar effects both in humans and

in macaques. The possible similarity in the acoustic cues of different affective states in humans and monkeys is treated in more detail elsewhere [3].

The present results suggest that there is a common reference / interpreting scheme with regard to the effects of emotional states on vocal behaviour, according to which the humans interpret all the animal vocalizations they encounter.

Listeners who had daily contact with animals as pets gave more correct interpretations than the others, which suggests that part of the ability to comprehend another species is acquired by experience. However, a proportion of correct responses well above chance level was reached even by those subjects who did not have daily contact with animals. This proves that such close contacts are not necessary for a certain ability to interpret correctly the sounds of another primate species.

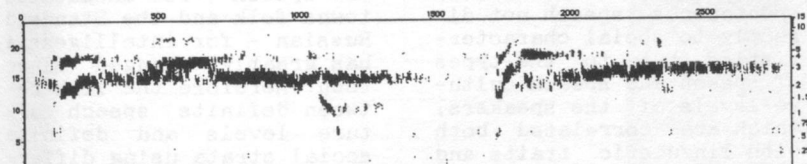
Evidence from this study lends support for the hypothesis that there is a common basis for the recognition of vocalizations between primate species.

6. REFERENCES

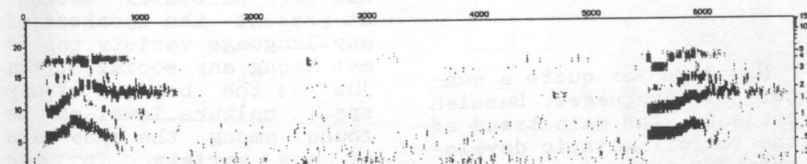
- [1] BERTRAND, M. (1969), "The behavioral repertoire of the Stumptail Macaque", Basel: Karger.
- [2] CHEVALIER-SKOLNIKOFF, S. (1974), "The ontogeny of communication in the Stumptail Macaque (*Macaca arctoides*)", Basel: Karger.
- [3] LEINONEN, L., LINNANKOSKI, I., AULANKO, R. & LAAKSO, M. (in preparation), "Interspecies vocal communication of emotions: Man and Macaque."
- [4] SCHERER, K.R. (1981), "Speech and emotional states". In *Speech Evaluation in Psychiatry* (J.K. Darby, editor), 189-220. New York: Grune & Stratton.
- [5] TOIVONEN, R. (1988) Intelligent Speech Analyser (ISA) järjestelmän käyttöohje. Tampere, 22.6.1988.
- [6] TOIVONEN, R. (1990) Tiivistetty ISAn käyttöohje. Tampere, 13.6.1990.
- [7] WILLIAMS, C.E. & STEVENS, K.N. (1981), "Vocal correlates of emotional states". In *Speech Evaluation in Psychiatry* (J.K. Darby, editor), 221-240. New York: Grune & Stratton.

TABLE 1. Percentages of human interpretations of vocalizations representing different emotional-motivational states of stumptail macaques. (Proportions less than 10 % are not indicated in the table. The number of vocalizations representing each behavioural state in the test material is given in parenthesis.) The figures on the diagonal indicate the proportion of responses considered "correct" in the strictest sense.

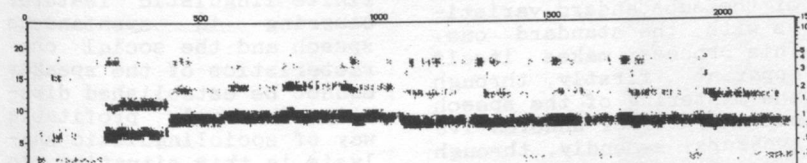
Behavioural states	Human interpretations						
	angry	frightened	ecstatic	commanding	submissive	satisfied	calling
aggression (4)	29	17	22	14			13
fear (2)	24	60					
sexual arousal (2)			46	32			
dominance (1)	11			84			
submission (4)					50	15	23
contentment (3)				10		68	12
contacting (2)		18			12		50



A. Female screams of fear



B. Female contact call



C. Female cooing

FIGURE 1. Examples of the macaque vocalizations analyzed with a psychoacoustic spectrogram program, a specialty of the Intelligent Speech Analyser system [5, 6]. Horizontal axis = time (ms), vertical axis = frequency in Bark (left) and kHz (right).