

Brain, Behavior and Evolution

Editor-in-Chief: W. Riss, Brooklyn, N.Y.

Publishers: S. Karger, Basel

Reprint (Printed in Switzerland)

Brain Behav. Evol. 16: 443-460 (1979)

Perception of Conspecific Vocalizations by Japanese Macaques

Evidence for Selective Attention and Neural Lateralization

Michael D. Beecher, Michael R. Petersen, Stephen R. Zoloth,

David B. Moody and William C. Stebbins

Kresge Hearing Research Institute, University of Michigan, Ann Arbor, Mich.,
and Center for Field Research, Rockefeller University, Millbrook, N.Y.

Key Words. Vocal communication · Animal communication · Neural lateralization · Selective attention · Animal psychophysics · Hearing · Primates

Abstract. Japanese macaques (*Macaca fuscata*) and control species (vervet, pigtailed macaque, bonnet macaque) were trained for food to respond to one class of recorded fuscata vocalizations and to not respond to a second class. A measure of neural lateralization was obtained by presenting the stimuli randomly to the right or the left ear, and comparing performance in the two ears (ear advantage method). Vocalizations were from Steven Green's field tapes. In experiment I, the two classes were Green's 'smooth early high coos' (SE) and 'smooth late high coos' (SL). Experiment II utilized the same vocalizations, but sorted into a high-pitched and a low-pitched class, i.e., orthogonally to the communication-relevant dimension. We found that (a) Japanese macaques learned the SE-SL discrimination faster than the pitch discrimination; (b) the reverse was true for the controls; (c) Japanese macaques showed a right-ear advantage (presumed left hemisphere advantage) for the SE-SL distinction, but not for the pitch discrimination, and (d) controls (with one exception) showed no ear advantage for either discrimination. These demonstrations of selective attention to communication-relevant parameters of conspecific vocalizations, and neural lateralization in the perception of these vocalizations, parallel similar findings in human speech perception.

Evidence that higher vertebrates utilize specialized processing mechanisms in the analysis of their acoustic communication signals is, so far, mostly limited to human speech and bird song. The bird studies, though they approach the problem from the motor rather than the sensory side, suggest sensory-perceptual 'templates' as the basis for selective vocal learning [Marler and Mundinger, 1971; Marler and Peters, 1977]. While recent evidence from speech perception studies has not in general favored the concept of species- and speech-specific processors at the sensory level

[Cutring, 1976; Kuhl, 1979; Pisoni, 1979], the human clinical evidence suggests that such mechanisms exist at a higher, cortical level [Kimmura, 1975]. There are intriguing parallels between the perception of speech by humans and song by birds, as has been pointed out by Marler [1970a, b, 1975]. Thus, it would undoubtedly be fruitful to extend the analysis of these sorts of phenomena to other animals, with nonhuman primates being particularly interesting candidates. Studies of primate perception of vocal communication signals will assist us in testing theories concerning the evolution of acoustic communication systems [Beecher, 1975; Marler, 1975; Zoloth and Green, 1979].

We are studying the perception by monkeys of conspecific vocalizations, using the methods of animals psychophysics [Stebbins, 1970]. We believe that there are three criteria that a behavioral approach must satisfy to provide evidence that a particular species utilizes specialized mechanisms for the analysis of conspecific acoustic signals. First, it must be shown that the study species processes a communication-relevant acoustic distinction more readily than do control species (which do not have the same distinction in their vocal repertoire). Second, control experiments are required to establish that the species difference is truly specific to the communication distinction. That is, generalized learning or hearing super-riorities have to be ruled out. Third, the nature of the presumptive mechanism must be characterized in some way. For example, categorical perception and neural lateralization are two key characteristics of human speech perception (though not necessarily unique to it). The critical point is to show that the study species is analyzing the communication distinction in a different way than the control species. Ultimately, as a fourth objective, *how* the mechanism enables more efficient processing should be demonstrated. If these behavioral experiments strongly suggest the existence of specialized mechanisms, then we can turn to physiological and neuropsychological methods to directly explore these mechanisms, and to developmental studies to examine their ontogeny. Throughout this paper we use the term 'special processing mechanism' simply to connote that the three above criteria have been fulfilled, and we hope the reader will regard it in this restricted sense. We specifically *exclude* these surplus connotations: that the mechanism is 'innate', genetically-specified or independent of experience. Our behavioral studies would logically be complemented by well-designed developmental studies which analyze the relative roles of genetic mechanisms and experience in the ontogeny of such processing mechanisms [Marler, 1970; Morse, 1979; Kuhl, 1979] (also

see 'Discussion'). The term 'strategy' could be used interchangeably with 'mechanism'; the point is that logically there must be a neural substrate for this mechanism/strategy, and though with these experiments we cannot specify its origin or its location, we can build a case for its existence.

The research described in this paper is directed toward satisfying, or rejecting, the three basic criteria for our study species, the Japanese macaque (*Macaca fuscata*). We give here a preliminary report; a more detailed presentation is in preparation. The Japanese macaque's vocal repertoire has been analyzed by Green [1975; see also Zoloth and Green, 1979]. At present, Green's description of this vocal communication system is the most complete available for any primate species. Our first experiments have utilized two subclasses of the macaque 'coo' vocalization. All macaque species evidently give the 'coo' vocalization, and in *M. fuscata*, Green recognizes seven subtypes, two of which he refers to as smooth-early-high coos (SE) and smooth-late-high coos (SL). SES are contact coos, typically given by isolated animals, by animals in small groups away from the main troop, or by young animals within the troop but away from relatives or playmates. They are often responded to antiphonally, and appear to function to maintain group cohesion. The animal giving an SE usually appears to be in a relatively calm state.

SLS may also be regarded as contact calls but differ in that they are active solicitations. For example, they are used in sexual solicitation by estrus females in the early stages of consortship. The call is typically given by a subordinate animal to a dominant animal. An animal giving an SL generally is in a somewhat more aroused state than an animal giving an SE. A detailed analysis of the context in which these and the other coo-subtypes occur can be found in Green [1975] and Zoloth and Green [1979]. Examples of the two call classes are shown in figure 1. SES and SLS contain a single, smooth frequency sweep, and they can be characterized by the position of the 'peak', or point of highest frequency: in SES the peak occurs in the first two-thirds of the call, in the SLS in the final third. Thus, examples of these calls can be conceived of as falling along a continuum of 'peak position'.

Experiment I: Peak-Relevant, Pitch-Irrelevant Discrimination

Our behavioral procedure is a variant of that which we routinely use to investigate acoustic difference thresholds [Moody *et al.*, 1976]. We

modified it in experiment I to compare the ability of Japanese macaques and other ('control') species to discriminate the class SE from the class SL; since many different acoustic variants make up a particular class, the task requires concept formation in addition to sensory discrimination. The stimuli were all taken from Green's field tapes, and therefore we can specify for each stimulus not only its acoustic parameters but also its field context. It can be seen from figure 1 that there is considerable variation in dimensions other than peak position. The set was explicitly balanced with respect to frequency (mean starting frequency for the SE set is 648 Hz with range 525-750 Hz, for the SL set 610 Hz with range 495-750 Hz). It can be seen that duration, extent of modulation and other variables also vary across the SE-SL classification. Thus, the task required that the animal attend to the relevant dimension while ignoring irrelevant variation occurring along other dimensions of the calls.

The animal was tested in a primate chair in an acoustic chamber, and the stimuli were presented monaurally through earphones. The response device was a metal, contact-sensitive tube which housed a small lamp. When the light was flashing the animal could initiate stimulus presentation by placing its hand on the tube. The light then became steady and the animal heard the background or negative stimulus repeating at a rate of once per 3 sec. Stimuli were delivered to either the left or right ear, according to a random order. This allowed a measure of ear advantage or laterality (discussed below). Wide-band, low-level noise was presented to both ears to mask the background noise which overlapped the field-recorded signals. The negative stimulus in experiment I was the SL class: in a given session anywhere for 1-8 different tokens made up the class. After a variable number of SLs (range 1-5, mean 3), an SE was inserted into the series. If the monkey released the tube within 2 sec of the SE, the response was defined correct and reinforced with a banana pellet. Two sorts of errors were possible. The animal could fail to release to a positive stimulus, i.e., fail to call an SE an SE, or could release to a negative stimulus (a 'false alarm'), i.e., call an SL an SE. 'False alarms' produced a 5-sec 'time-out' (signalled by the tube light going out), during which further responses had no effect other than resetting the 5-sec time-out.

The animals were first trained to discriminate a single SE from a single SL (the first pair in fig. 1). Once they had reached criterion (defined below) on this '1-1' set, they were asked to discriminate two tokens of each class (the first two pairs in fig. 1). The '2-2' set was the first stage calling for a conceptual distinction on the monkey's part, by requiring at-

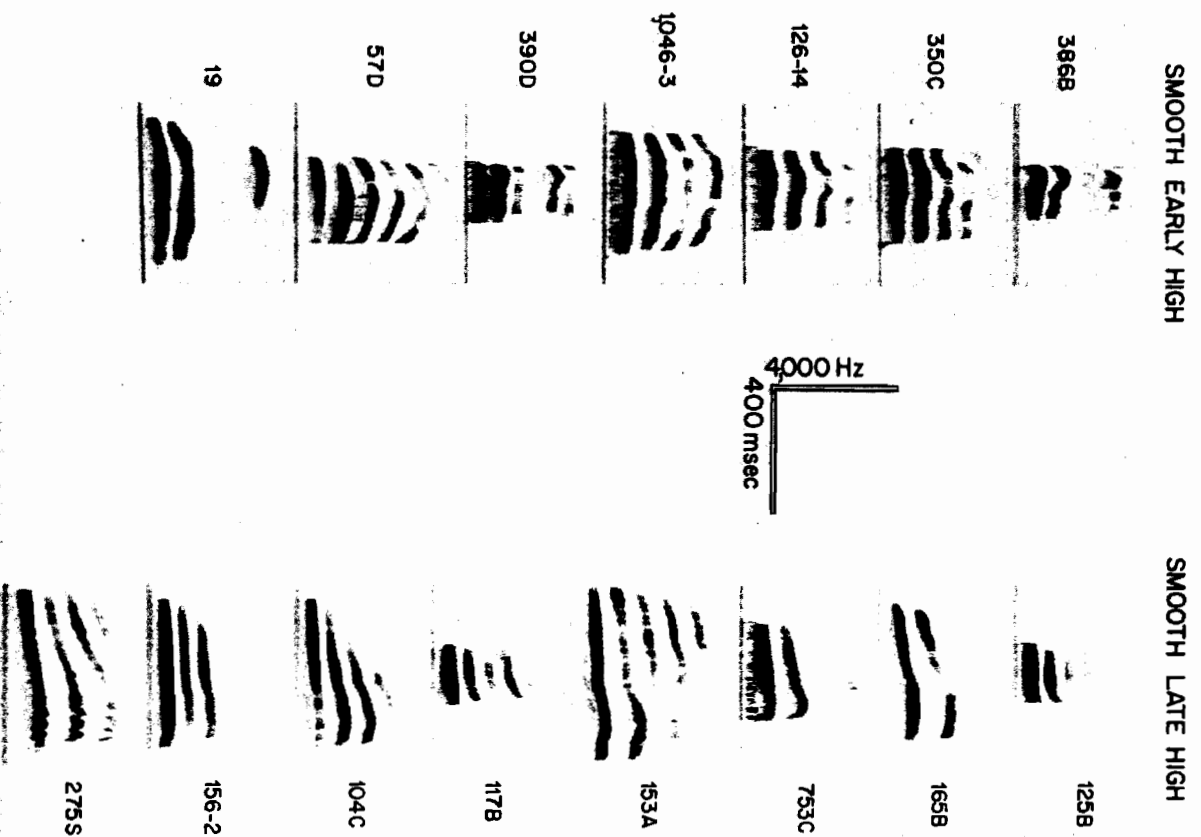


Fig. 1. Sonagrams of smooth early highs and smooth late highs. Given in the order they were introduced. The defining characteristic is the occurrence of the frequency peak early or late in the call. Note variability in other acoustic parameters, including frequency, duration, harmonic emphasis, and rate of frequency modulation.

tention to the relevant dimension in preference to irrelevant ones. Thereafter, when the animal reached criterion on a particular stage, the number of tokens in each class was increased. The maximum number used in these experiments was 7 SEs and 8 SLs. The three criteria for transfer to the next more difficult set were (1) two consecutive sessions of at least 90% correct responses (releases to SEs) overall; (2) better than 80% for each particular SE, and (3) less than 10% false alarms (releases to SLs).

The subjects for experiment I were 3 Japanese macaques and 3 control animals: a pigtailed macaque (*Macaca nemestrina*), a bonnet macaque (*Macaca radiata*) and a vervet monkey (*Cercopithecus aethiops*). The two macaque control species also make 'coo' sounds, but there have been no field studies of their vocal communication systems, so we do not know if they have a distinction corresponding to SEs and SLs. The vervet does not produce a coo vocalization [Srnhsaker, 1967].

We describe the discrimination in experiment I as 'peak-relevant, pitch-irrelevant' since we systematically matched the SE and SL sets for frequency: the SEs averaged slightly higher with respect to starting frequency (75 Hz) but slightly lower with respect to average frequency (25 Hz). This match is about as good as one can obtain with natural tokens. The stimuli were introduced in the order shown in figure 1. The number of sessions each animal required to reach criterion for each successive stimulus set is shown in figure 2. Clearly, the task was easier for the Japanese macaques than in was for the controls. The 2-2 set, the first one requiring a conceptual distinction, was particularly difficult for the control animals, one (M88) who was unable to reach criterion on this set, and a second (M58) who finally did but failed to on the 3-3 set. These 2 animals were subsequently switched to a procedure in which the stimuli were introduced in a different order, and ultimately reached criterion on the 8-7 set; these data are not included in figure 2 since they are not comparable. Thus, though all animals ultimately mastered the discrimination, the control animals found it a much more difficult task.

Experiment II: Pitch-Relevant, Peak-Irrelevant Discrimination

Experiment II utilized the stimuli of experiment I, but resorted to produce a pitch-relevant, peak-irrelevant discrimination: the experiment was designed to determine whether the species difference of experiment I was in fact due to the particular communicative significance of peak posi-

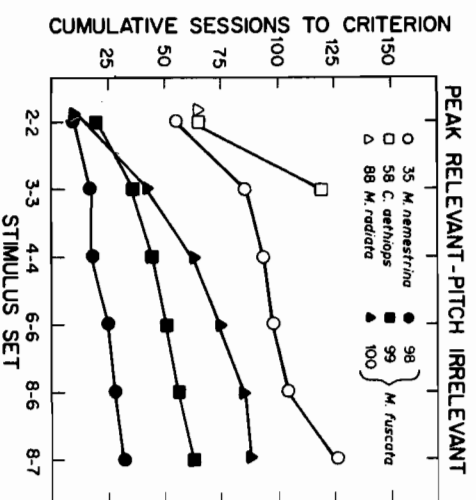


Fig. 2. Experiment I. Cumulative sessions to criterion for the 6 animals as a function of the size of the stimulus set. '8-7' (for example) means that there were 8 SLs and 7 SEs in the set (the full set in fig. 1), '4-4' that there were 4 of each (the top 4 in each column of fig. 1). Thus, the horizontal dimension represents increasing complexity or difficulty of the discrimination. M58 was unable to reach criterion at 3-3, M88 at 2-2; they were switched to another, easier version of the discrimination (not shown).

tion for the Japanese macaque and the lack of same for the control animals. If peak position is indeed an inherently relevant parameter for the Japanese macaque, then these animals should have difficulty with a task in which peak position is irrelevant, i.e., conflicts with the rewarded (relevant) cue. This expectation would not hold for the control animals; they indeed might find the pitch cue more salient than the peak position cue. This design is analogous to that of experiments on speech perception which use frequency as a nonlinguistic dimension [Kuhl, 1976; Springer, 1973; Wood *et al.*, 1971], though the communicative significance of pitch for Japanese macaques (and the control species) is unclear.

Two new classes were formed using 12 of the 15 stimuli of experiment I: a high-pitched class (mean starting frequency 710 Hz and range 675-750 Hz) and a low-pitched class (mean starting frequency 518 Hz and range 495-550 Hz). Each of these two pitch classes included several SEs and several SLs. Hence, pitch was a consistent cue while peak was an 'irrelevant' or distracting cue. The low-pitched stimuli were arbitrarily designated the negative class, the high-pitched stimuli the positive class. An-

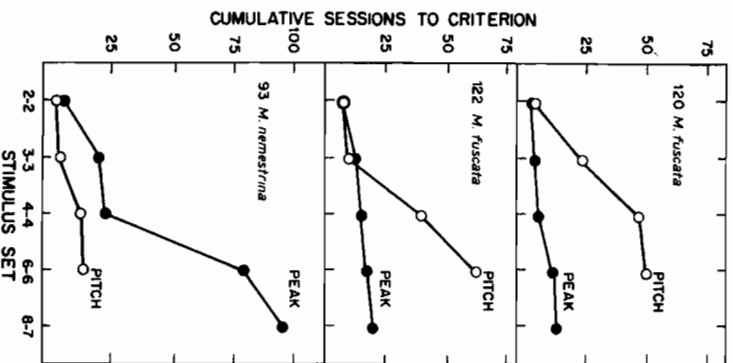


Fig. 3. Experiment II. Cumulative sessions to criterion as a function of size of stimulus set. The stimuli on the 'peak-relevant' discrimination are the same as in experiment I, but were introduced in an order that made the discrimination less difficult. The 'pitch relevant' discrimination utilized 12 of the 15 stimuli.

alogous to experiment I, a variable number of low-pitched stimuli preceded each presentation of a high-pitched stimulus, and a release within 2 sec of the high-pitched stimulus was reinforced.

4 new animals were used (2 Japanese and 2 control). Performance on this 'pitch-relevant, peak-irrelevant' discrimination was compared with that on a 'peak-relevant, pitch-irrelevant' discrimination where the same 12 stimuli (and three additional stimuli) were sorted into the appropriate peak position categories. The order of exposure to each discrimination was counterbalanced for the 4 animals. The data are presently incomplete for 1 animal, but the results for the other 3 are summarized in figure 3. It can be seen that the 2 Japanese animals learned the peak-relevant discrimi-

mination more readily than the pitch-relevant discrimination. The control animal (a pigtailed macaque), on the other hand, learned the pitch-relevant discrimination more quickly than the peak-relevant discrimination. Note also that the 2 Japanese animals learned the peak discrimination faster than did the control animal (replicating experiment I) but that the control animal learned the pitch discrimination faster than did the 2 Japanese animals. The preliminary results, therefore, are consistent with our interpretation of the previous experiment: in comparison to the control species, Japanese macaques are predisposed to acquire a discrimination that is based on a natural classification of their own communication system.

We believe experiment II eliminates a number of general explanations of the species difference, including superior sensory capacities or problem-solving ability in the Japanese macaques. The interpretation we favor is that of selective attention: the Japanese macaques naturally are more attentive to the peak dimension than are the comparison species. This finding is akin to that, for example, of *Kuhl* [1976, 1979] who found that human infants are more attentive to vowel category than pitch contour. In any case, our experiments only suggest a species difference in processing mechanism of strategy; this interpretation must be bolstered by direct evidence that the Japanese macaques are indeed processing the stimuli in some fundamentally different way from the control animals. This sort of evidence, which may allow us to characterize the presumed mechanism, we have attempted to gather in additional experiments. In this paper we will consider only one area: neural lateralization.

Neural Lateralization: Ear Advantage Measurements

One indication of left hemisphere specialization for speech processing in intact humans is the ubiquitous, though small, performance advantage enjoyed by the right ear for the identification of speech sounds [see review in *Kimura, 1975; Catlin et al., 1976*]. Not only is a right ear advantage (REA) typically found for speech stimuli, but conversely, nonspeech stimuli (e.g., music) generally give a left-ear advantage (LEA) [*Kimura, 1975*]. Although there is not complete agreement as to the neural mechanism underlying the ear-advantage data, there is a clear correlation between laterality as measured by ear advantage and laterality of hemispheric processing as determined by independent (and more direct) meth-

ods: speech is processed (typically) by the left hemisphere and gives an REA, and nonspeech acoustic stimuli are processed predominantly in the right hemisphere and give an LEA.

Given the evolutionary relationship of human and other primates, it is possible that analyzing mechanisms for communication in monkeys are also lateralized in one hemisphere, the logical guess being the left. As indicated earlier, we obtained a measure of laterality in these experiments by presenting the stimuli randomly to the left and right ears. The two experiments present a powerful test for laterality since we can compare performance, using the same stimuli, on a discrimination in which peak position is the relevant cue with one in which it is an irrelevant one. This design is similar to that of *Wood et al.* [1971]: using an evoked-potential method they found evidence of greater left-hemisphere activity on a task requiring classification of speech sounds according to phoneme category but no hemisphere difference on a task requiring classification of the same sounds according to pitch.

Two problems encountered by nearly all human auditory laterality studies are (a) subjects perform nearly perfectly, producing few errors with either ear, and (b) partially as a function of (a), the performance differences between the two ears are small. Since these problems are especially evident when monaural listening procedure is used, the dichotic listening procedure is generally the method of choice. We selected the monaural listening task, however, because it would simultaneously yield information on selective attention and laterality.

The ear advantage data are summarized in table I. Laterality was measured by comparing left-ear and right-ear performance for each stimulus (SE) for each session over the entirety of experiments I and II. The measure of REA is simply the proportion (p) of those cases in which right-ear performance was better than left-ear performance (ties were arbitrarily distributed half into each category). Thus, if measured ear advantage fluctuates randomly, p should be approximately 0.5, an REA would be indicated by a value significantly greater than 0.5, an LEA by one significantly smaller than 0.5.

There are two major findings, corresponding to the two experiments. (1) All 5 Japanese macaques showed statistically significant REAs on the peak-relevant task. However, only 1 of the 4 control animals (the vervet) showed a significant ear advantage (it was also an REA). (2) The 2 Japanese macaques tested on both the pitch-relevant task and the peak-relevant task (experiment II), showed a significant difference in p for the two

Table I. Proportion of cases in which right ear performance is better than left ear performance

Subject	Species	Peak-relevant	Pitch-relevant	Difference
M35	<i>M. nemestrina</i>	0.52		
M58	<i>C. aethiops</i>	0.61*		
M88	<i>M. radata</i>	0.53		
M93	<i>M. nemestrina</i>	0.51	0.53	-0.02
M98	<i>M. fuscata</i>	0.59*		
M99	<i>M. fuscata</i>	0.73*		
M100	<i>M. fuscata</i>	0.61*		
M120	<i>M. fuscata</i>	0.62*	0.40*	0.22*
M122	<i>M. fuscata</i>	0.61*	0.51	0.10*

* $p < 0.05$ (z-test).

tasks, with significant REAs on the peak-relevant task and LEAs (significant in one animal but not the other) on the pitch-relevant task. The control animal showed no significant difference in p on the two tasks; interestingly, though, the ear advantage was opposite that shown by the 2 Japanese animals. Thus, according to the conventional interpretation of ear advantage data, we have demonstrated left hemisphere processing of conspecific vocalizations by Japanese macaques. It clearly is a rather *specific* effect, in that it was seen only in the Japanese animals (the vervet excepted) and then only when the task required classification according to peak position.

Discussion

The Present Research: Summary and Prospects

We suggested in the 'Introduction' that specialized processing mechanisms become a reasonable hypothesis when the investigators have (1) demonstrated more efficient processing of the communication distinction by the study species than by the control species; (2) provided evidence that different mechanisms are used by study species and control species, and (3) established by appropriate control experiments the specificity of both effects to the communication distinction. We believe that the data

presented above satisfy these criteria. Japanese macaques made the SE-SL distinction more readily than control species. They evidently processed the stimuli differently, since they lateralized the sounds while the controls (with one exception) did not. And finally, both effects were specific to the peak-relevant task and in fact were reversed on the peak-irrelevant task. It is tempting to suggest that the key condition for both the selective attention and the lateralization effects was the communicative relevance of the peak-position distinction for Japanese macaques. If this hypothesis is correct, we have a striking parallel with human speech perception. We will consider the implications of the neural lateralization data in more detail elsewhere [in preparation]. We should make several cautionary remarks here, however. First, the lateralization effect may not be attributable to the communicative relevance *per se* of peak position, but rather to some general property of the two discriminations (e.g., one is a temporal discrimination, the other a pitch discrimination). Humans do show REAs for some nonspeech stimuli [e.g., *Bever and Chiarello, 1974; Halperin et al., 1973; Papcun et al., 1974; Gordon, 1975; Blechner, in press*]. Second, *Green's* analysis of Japanese macaque coos did not partition SEs and SLs according to pitch, and it cannot be said that this dimension does not have communicative relevance; in fact, *Green* does use pitch to differentiate other coo classes. Moreover, the functional significance of peak position and pitch for the comparison species is simply unknown. Consequently, the discriminations cannot be neatly placed into communication-relevant and communication-irrelevant categories for these species, analogous to the phonetic and non-phonetic dimensions of human speech perception studies. These cautions notwithstanding, the species difference we have found still represents the first demonstration that nonhuman primates employ neurally lateralized mechanisms, presumably located in the left hemisphere, for the analysis of conspecific communication sounds.

If indeed Japanese macaques use specialized mechanisms in processing the SE-SL distinction, how should further analysis proceed? We are presently doing additional experiments to characterize the way in which the stimuli are perceived. (1) We are examining how the animals classify the stimuli besides peak position are relevant, and whether Japanese animals and controls differ with respect to the features they attend to. (2) We are examining how the animals respond to stimuli arrayed along a synthetic SE-SL dimension. In preliminary experiments, we have obtained conventional generalization gradients which indicate that peak position is indeed

a relevant, controlling dimension when all other dimensions (frequency, duration, etc.) are held constant. (3) We intend to examine categorical perception along this continuum. A necessary condition for the demonstration of categorical perception is a difference in discriminability along the continuum, with best discrimination occurring at the 'category boundary' [see *Kuhl, 1979, or Pisoni, 1979*]. Thus, these experiments will be designed to measure discrimination performance at points along the SE-SL continuum. (4) Other experiments may allow us to say whether these hypothetical mechanisms should be conceived of as basically sensory or as higher-level mechanisms.

The control experiments - necessary for demonstrating that the species difference is restricted to the communication context - are critical to our argument. We have been unable so far to perform what is perhaps the ideal control experiment. If adequate field tapes are available for two species, then acoustic distinctions unique to each repertoire could be used and the two species could exchange roles as study and control species. That is, when the distinction is drawn from species A's repertoire, species B is the control species, and when from species B's repertoire, species A is the control species.

When the presumptive mechanisms are well characterized, further research of a very different sort will be required. First, the developmental basis of this species difference should be analyzed: does it depend on critical acoustic experiences which were enjoyed by the Japanese macaques but not the control species before their arrival in the laboratory? Second, we reiterate that mechanisms can only be inferred from behavioral experiments. These experiments can direct our thinking about the nature of these mechanisms, but ultimately, other, more direct methods are needed to reveal their neural basis (physiological methods, e.g., *Newman and Wollberg [1973]*, or neuropsychological methods, e.g., *Dewson [1976]*).

Other Comparative Studies

The notion that humans utilize a special set of perceptual mechanisms for analyzing speech (the 'speech mode', *Liberman et al. [1967]*) has inspired in recent years a number of studies of speech perception by animals. Such comparative studies are the most direct way of determining whether speech sound analysis requires a species-specific speech mode, or can be carried out by generalized auditory processors. There have been a number of studies now which unequivocally demonstrate that a variety of

mammalian species can discriminate human speech sounds, the only limitations evidentially being due to peripheral sensory differences [see reviews by Kuhl, 1979; Miller, 1977]. The key question, however, is whether animals process these stimuli in similar or different ways than do humans. Research has focused on categorical perception, since this phenomenon was originally thought to be unique to the speech mode [Mattingly *et al.*, 1971]. Kuhl and Miller [1975] and Kuhl [1979] have shown that chinillas categorically perceive the voiced-voiceless distinction: their /b-/p/, /d-/t/, and /g-/k/ boundaries are virtually identical to those of humans and their ability to discriminate pairs along the /d-/t/ continuum is best at the boundary. Waters and Wilson [1976] and Morse and Snowdon [1975] also have found categorical perception in monkeys (i.e., better between-category than within-category discrimination), though the performance of monkeys and humans was not identical in all respects. The first clear demonstration of a species difference with respect to speech was obtained in our laboratory [Simoni *et al.*, 1976]. We found that humans showed an abrupt change in processing speed at the /ba-/da/ category boundary, taking longer to process within-category than between-category discriminations. Monkeys, on the other hand, showed no such abrupt change; processing speed changed uniformly with discrimination difficulty. The discrepancy in these studies may have a simple explanation. Methods or measures designed to assess discriminative capacity have so far shown no basic difference between humans and animals; it is possible that species differences will be revealed, however, when higher-level measures are used, such as response latency.

There is one recent study of categorical perception of animal communication sounds by an animal species. Snowdon [1979] presents data suggesting that marmosets categorically perceive a synthetic continuum constructed between models of marmoset 'open mouth trills' and 'closed mouth trills'; this dimension is constructed by varying a single parameter, duration. Humans, on the other hand, perceived this dimension continuously. This species difference must be regarded as tentative, however, since (a) no discrimination data were obtained, only labeling data, and thus, by the usual criteria, categorical perception cannot actually be said to have been demonstrated, and (b) very different procedures were used to assess perception in the two species, a psychophysical procedure for the humans, and an evoked vocal response, obtained from the colony as a whole, for the marmosets. Nevertheless, if the latter objection can be overcome, the data would suggest a processing difference at some level.

Need for Developmental Studies

The species differences we observed in the present study may ultimately be shown to depend upon a specific sort of developmental history. For example, in the simplest possible case, the species differences might be eliminated by exposing controls to Japanese macaque sounds at some point during development (undoubtedly an easier experiment to talk about than to do) or by depriving Japanese macaques of experience with these sounds. Other more complicated cases are possible, such as, for example, the selective learning Marler has found in white-crowned sparrows or swamp sparrows [Marler, 1970a, b; Marler and Peters, 1977]. To date, there have been only two studies of which we are aware on the ontogeny of communication signals in nonhuman primates. Winter *et al.* [1973] found that neither isolation nor deafening interfered with the production of the full vocal repertoire by squirrel monkeys as adults. Newman and Symmes [1974], on the other hand, found abnormalities in the clear calls (coos) of rhesus monkeys raised in partial social isolation; this is of interest since we are using homologous calls in the experiments described in this paper.

Developmental studies can serve an additional function: they can themselves provide indirect evidence for species-specific mechanisms. In fact, the developmental approach has provided some of the most compelling evidence to date for species-specific mechanisms in speech perception, certainly more compelling than the comparative evidence reviewed above. Cross-culture studies have revealed that individuals of language groups that do not have the prevoiced versus voiced or voiceless distinction, cannot hear this distinction, at least not in the experimental situation [Abramson and Lisker, 1972; Lasky *et al.*, 1975]. Similarly, Japanese do not make the /r-/l/ distinction and cannot in fact hear it [Miyawaki *et al.*, 1975]. Assuming that these differences are not linked in some trivial way to the methodologies used, they can be interpreted as indicating (1) the existence of different genetically-determined feature-detectors in the two groups; (2) the loss of same in the group that has not had the appropriate experience for maintaining them, or (3) the development of feature detectors through experiences that are appropriate to the phonetic environment of each group, i.e., whether or not the /r-/l/ distinction is made. We can eliminate the first alternative since infants and Japanese raised speaking English (or another language in which the /r-/l/ distinction is made) can hear the distinction [Eimas, 1975; Miyawaki *et al.*, 1975]. We cannot distinguish at present between the second and third alternatives [for a review of this literature see papers by Morse, 1979; Kuhl, 1979].

Final Remarks

This discussion has made no attempt to survey the evidence for special processing mechanisms for acoustic communication signals. Instead, we have sought to indicate the role behavioral experiments such as ours can have in this area of research. Optimally, these studies can build a strong, though circumstantial, case for mechanisms. A well-known example from vision is the use of phenomena such as the two-segment dark adaptation curve or the Purkinje shift as indirect evidence for the duplex retina. We believe we have presented here a clear, though preliminary, case for special processing mechanisms in Japanese macaques for the perception of their vocal communication signals. We hope to develop this case with further experiments.

References

- Abramson, A. S. and Lisker, L.: Voice-timing perception in Spanish word-initial stops. *Status Rep. Speech Perception, Haskins Lab. January-June: 15-25 (1972).*
- Beecher, M. D.: Evolution of hearing and acoustic communication in vertebrates. *J. acoust. Soc. Am. 58: S18 (1975).*
- Bever, T. and Chiarello, R.: Cerebral dominance in musicians and nonmusicians. *Science 185: 537-539 (1974).*
- Blechner, M.: Right-ear advantage for musical stimuli differing in rise-time. *Percept. Psychophys. (in press).*
- Catlin, S.; Van Derveer, N., and Teicher, R.: Monaural right-ear advantage in a target-identification task. *Brain Lang. 1976: 470-481.*
- Cutting, J. E.: There may be nothing peculiar to perceiving in a speech mode. *7th Int. Symp. Attention and Performance, Senanque 1976.*
- Dewson, J. H., III: Preliminary evidence of hemispheric asymmetry of auditory function in monkeys; in Harnad, Doty, Goldstein, Jaynes and Krauthamer, *Lateralization in the nervous system (Academic Press, New York 1976).*
- Eimas, P.: Auditory and phonetic coding of the cues for speech. *Discrimination of the /r-l/ distinction by young infants. Percept. Psychophys. 18: 341-347 (1975).*
- Gordon, H.: Hemispheric asymmetry and musical performance. *Science 189: 68-69 (1975).*
- Green, S.: Communication by a graded vocal system in Japanese monkeys; in Rosenblum, *Primate behavior, vol. 4, pp. 1-102 (Academic Press, New York 1975).*
- Halperin, Y.; Nachson, I., and Carmon, A.: Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli. *J. acoust. Soc. Am. 53: 46-50 (1973).*
- Kimura, D.: Cerebral dominance for speech; in Tower, *Human communication and its disorders, vol. 3, pp. 365-371 (Raven Press, New York 1975).*
- Perception of Conspecific Vocalizations by Japanese Macaques
- Kuhl, P. K.: Speech perception in early infancy; in Hirsch, Eldridge, and Silverman, *Hearing and Davis: Essays honoring Hallowell Davis (Washington University Press, St. Louis 1976).*
- Kuhl, P. K.: Recent findings in speech perception. The case for speech-sound categorization by general auditory mechanisms. *Brain Behav. Evol. (in press).*
- Kuhl, P. K. and Miller, J. D.: Speech perception by the chinchilla. Voiced-voiceless distinction in alveolar plosive consonants. *Science 190: 69-72 (1975).*
- Lasky, R.; Syrdal-Lasky, A., and Klein, R.: VOT discrimination by four six month-old infants from Spanish environments. *J. exp. Child Psychol. 20: 215-225 (1975).*
- Liberman, A. M.; Cooper, F. S.; Shankweiler, D. S., and Studdert-Kennedy, M.: Perception of the speech code. *Psychol. Rev. 74: 431-461 (1967).*
- Marler, P.: A comparative approach to vocal development. Song learning in the white-crowned sparrow. *J. comp. Physiol. Psychol. 71: 1-25 (1970a).*
- Marler, P.: Birdsong and speech development. Could there be parallels? *Am. Scient. 58: 669-673 (1970b).*
- Marler, P.: On the origin of speech from animal sounds; in Kavanagh and Cutting, *The role of speech in language (MIT Press, Cambridge 1975).*
- Marler, P. and Mundinger, P.: Vocal learning in birds; in Moltz, *Ontogeny of vertebrate behavior, pp. 389-450 (Academic Press, New York 1971).*
- Marler, P. and Peters, S.: Selective vocal learning in a sparrow. *Science 198: 519-521 (1977).*
- Mattingly, I. G.; Liberman, A. M.; Syrdal, A., and Hawkes, T.: Discrimination in speech and nonspeech modes. *Cogn. Psychol. 2: 131-157 (1971).*
- Miller, J. D.: Perception of speech sounds in animals. Evidence for speech processing by mammalian auditory mechanisms; in Bullock, *Recognition of complex acoustic signals, pp. 49-58 (Verlagsgesellschaft Berlin, 1977).*
- Miyawaki, K.; Strange, W.; Verbrugge, R.; Liberman, A. M.; Jenkins, J. J., and Fijimara, O.: An effect of linguistic experience. The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Percept. Psychophys. 18: 331-340 (1975).*
- Moody, D. B.; Beecher, M. D., and Stebbins, W. C.: Behavioral methods in auditory research; in Smith and Vernon *Handbook of auditory and vestibular research (Thomas, Springfield 1976).*
- Morse, P. A.: The infancy of infant speech perception: the first decade of research. *Brain Behav. Evol. 16: 351-373 (1979).*
- Morse, P. A. and Snowdon, C. T.: An investigation of categorical speech discrimination by rhesus monkeys. *Percept. Psychophys. 17: 9-16 (1975).*
- Newman, J. D. and Symmes, D.: Vocal pathology in socially deprived monkeys. *Dev. Psychobiol. 7: 351-358 (1974).*
- Newman, J. D. and Wollberg, Z.: Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. *Brain Res. 54: 287-304 (1973).*
- Nottebohm, F.: Vocal behavior in birds; in Farner, *Avian biology, vol. V (Academic Press, New York 1975).*
- Pisoni, D. B.: On the perception of speech sounds as biologically significant signals. *Brain Behav. Evol. 16: 330-350 (1979).*

- Papcun, G.; Krashen, S.; Terbeek, D.; Remington, R., and Harszman, R.: Is the left hemisphere specialized for speech, language and/or something else? *J. acoust. Soc. Am.* 55: 319-327 (1974).
- Sinnott, J. M.; Beecher, M. D.; Moody, D. B., and Stebbins, W. C.: Speech sound discrimination by monkeys and humans. *J. acoust. Soc. Am.* 60: 687-695 (1976).
- Snowdon, C. T.: The response of non-human animals to speech and to species-specific sounds. *Brain Behav. Evol.* 16: 409-429 (1979).
- Springer, S. P.: Memory for linguistic and nonlinguistic dimensions of the same acoustic stimulus. *J. exp. Psychol.* 101: 159-163 (1973).
- Stebbins, W. C.: Animal psychophysics: The design and conduct of sensory experiments (Appleton-Century-Crofts, New York 1970).
- Srulsaker, T.: Auditory communication among vervet monkeys (*Cercopithecus aethiops*); in Altman, Social communication among primates, pp. 281-324 (University of Chicago Press, Chicago 1967).
- Waters, R. A. and Wilson, W. A., jr.: Speech perception by rhesus monkeys. The voicing distinction in synthesized labial and velar stop consonants. *Percept. Psychophys.* 19: 285-289 (1976).
- Winter, P.; Handley, P.; Ploog, D., and Schott, D.: Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour* 47: 320-339 (1973).
- Wood, C. C.; Goff, W. R., and Day, W. S.: Auditory evoked potentials during speech perception. *Science* 173: 1248-1251 (1971).

Dr. Michael Beecher, Department of Psychology, University of Washington,
Seattle, WA 98195 (USA)

Supported by a grant from the
National Science Foundation
"B N S 77-19254"