International Journal of Comparative Psychology, Vol. 4, No. 1, 1990

THE PERCEPTION OF COMPLEX ACOUSTIC PATTERNS IN NOISE BY BLUE MONKEY (*CERCOPITHECUS MITIS*) AND HUMAN LISTENERS

Charles H. Brown
Joan M. Sinnott
University of South Alabama

ABSTRACT: Blue monkeys (*Cercopithecus mitis*) were trained to detect complex acoustic signals embedded in noise. Masked thresholds were determined for four human consonant-vowel speech sounds (*ba, pa, ga,* and *ka*), and four blue monkey calls (*boom, pyow, chirp,* and *trill*). The ability of monkey listeners to hear these signals in noise was compared with humans. Results showed that monkey and human hearing was very similar. The mean difference between species for these eight stimuli in the broad-band noise environment was 2.3 dB. The signal-to-noise ratio for perception ranged from 4.8 dB for the *ka* to -23.8 dB for the *boom*. The four monkey calls were audible at a signal-to-noise level that was 8.1 dB less than that required for the detection of the speech sounds. However, most of this effect was due to the audibility of the *boom*. With the *boom* excluded, the mean signal-to-noise ratio for detection of the remaining 7 sounds was -0.5 dB, and the mean difference in the audibility of the speech and monkey sounds within this set was 2.6 dB. These results contrast with previous findings which used simulated rain forest noise as the masking noise (Brown, 1986). In rain forest noise, test signals were audible at signal-to-noise ratios approximately 10 dB less than those reported here, and the observed difference in the relative audibility of human and monkey utterances was larger. These findings suggest that rather small variations in the amplitude and spectrum of the ambient noise may have a strong influence on the audibility of vocal signals in nature.

INTRODUCTION

Many different species of animals are conspicuously vocal and they may signal acoustically over long distances in relatively noisy habitats (Moynihan, 1966; Payne & Webb, 1971; Morton, 1975; Gautier & Gautier, 1977; Byrne, 1981; Brenowitz, 1982; Brown & Schwagmeyer, 1984; Ryan & Brenowitz, 1985; Waser & Brown, 1986). The long-range calls of some forest primates, for example, may be audible for distances approaching 2 km (Brown, 1989a), and both the coordination of movement within a social unit, and the regulation of spac-

---

Address correspondence to Dr. Charles H. Brown, Department of Psychology, University of South Alabama, Mobile, AL 36688, USA.

ing between competing social units appear to be dependent upon the successful exploitation of the acoustic modality. Listeners are confronted with the task of distinguishing species-specific vocalizations from the calls of sympatric species. For example, in many instances forest monkeys may live sympatrically with perhaps a half dozen other species of primates, with scores of species of birds, and with hundreds of species of insects. Because rain forests exhibit a diversity of niches arrayed vertically as well as horizontally, this habitat exhibits a multitude of biotic sources of acoustic signals, and this rich assemblage of organisms in concert may generate a cacophony of background noise as each species struggles to make its own message heard (Waser & Brown, 1986; Brown & Waser, 1988).

Though the rain forest habitat may represent an end point for acoustic competition among vocalizers, the problem of hearing in noise is by no means restricted to rain forest inhabitants. In the tropical zone, the residents of the riverine forests and the savanna, and in the temperate zone the inhabitants of the grass lands, the swamps and the forests all encounter similar difficulties. Furthermore, in many instances the noise generated by other organisms is overshadowed by the noise produced by rustling vegetation, rain and thunder. Hence, in many niches an evolutionary premium may have been placed on the ability of organisms to detect and classify complex acoustic wave forms embedded in ongoing temporally fluctuating background noise (Brown & Waser, 1988).

Natural selection could act to heighten the audibility of vocal signals in noise by “improving” the acoustical structure of the signal so that it is more perceptible in noise, by “improving” the structure of the signal detection module of the auditory system so that it can “hear” signals at lower signal-to-noise ratios, or by simultaneously “improving” the design features of both the sound production and sound reception systems. It is clear that different species employ species-specific communication sounds that differ substantially in their audibility in noise. Blue monkeys (*Cercopithecus mitis*) and grey-cheeked mangabeys (*Cercocebus albigena*) have been shown to be able to hear their own species utterances at signal-to-noise levels that are about 10 dB below the level at which humans can hear speech sounds (Brown, 1986). Yet, the acoustical or perceptual reasons for these observed differences in the audibility of various signals is not fully known. Brown (1986) suggested that the vocalizations of rain forest monkeys evolved a structure that rendered them less susceptible to masking than human speech. However, it is unclear if calls produced by some primates have acquired a structure which in general is inherently more audible in any noise, or if selection has acted to “adjust” the acoustical features of various primate repertoires to cope specifically with the acoustics of the local habitat. Hence, in this context, speech sounds may be found to be more audible than monkey calls if testing were conducted in an acoustic environment that corresponded to that in which speech sounds originally evolved.

In the present paper we explore this possibility by comparing the audibility of selected speech and monkey utterances in an acoustic environment which provides a spectral contrast to the rain forest habitat.

METHOD

Animals

Three adult male blue monkeys (*Cercopithecus mitis*) served as subjects. The monkeys were laboratory born, and ranged from five to six years in age. The primate laboratory animal care standards exceed both institutional and NIH guidelines. Experimental animals were weighed daily, and were maintained on a diet of Purina monkey chow supplemented with PRIMA-Treats and a diverse diet of fresh fruit and vegetables. Animals were fed twice daily, and maintained in a colony situation promoting social exchanges. Animals, individually housed, were supplied with foraging boxes, and had scheduled access to a "gymnasium" supplied with swings and other "toys." Attempts were made to breed selected animals. All monkey subjects were experimentally naive at the onset of training. Three human listeners (two males and one female) were also tested. The human subjects ranged in age from 21 to 28 years of age. All humans and monkeys were screened for hearing within normal limits for frequencies ranging from 125 Hz to 8,000 Hz.

Apparatus

Monkeys and humans were tested in a double-walled IAC model 1203 semi-anechoic room. The walls and ceiling of the room were covered with 10 cm thick Sonex acoustic foam. The absorption coefficient of this treatment was greater than 0.6 at 125 Hz and increased to 0.9 or more for frequencies above 400 Hz. The floor of the room was carpeted. Stimuli consisted of the human consonant-vowel speech sounds *ba, pa, ga*, and *ka*, and the blue monkey calls *pyow, boom, trill*, and *chirp*. This set of vocalizations samples the diverse range of calls in the blue monkey's repertoire (Marler, 1973). The *pyow* and *boom* calls are loud calls emitted exclusively by adult males. The *trill* and *chirp* calls are given by both juveniles and adult females. These vocalizations were selected because it is possible to obtain high-quality recordings of these signals in an acoustically controlled laboratory context. All eight test stimuli were recorded in the laboratory. The speech sounds were recorded from a male talker with a Standard American dialect. The *pyow* and *boom* monkey calls were recorded from an adult male blue monkey, and the *trill* and *chirp* calls were recorded from an adult female blue monkey. The test signals were digitized on a 12-bit A/D converter at a sample rate of 27.5 kHz. Monkeys were seated in a Primate Products monkey chair during the course of the one hour daily session. Humans were seated in a conventional institutional chair in the same position in the sound field occupied by the monkeys. Auditory stimuli were presented free-field through a Polk 10 loudspeaker, and stimulus presentations were controlled by a Compaq 486/25 computer. The test stimuli were embedded in a broad band masking noise produced by a General Radio 1381 noise generator. The masking noise was presented at a level of 60 dBA.

Procedure

During 1 h daily sessions monkeys were trained with operant conditioning positive reinforcement procedures to contact a response key to initiate a trial and to terminate key contact upon detecting the presentation of an acoustic signal. The presentation of the acoustic stimuli was controlled by computer and occurred unpredictably according to a variable interval schedule (1-7 s). Following sound presentation, the termination of key contact within a 1.5 s response interval resulted in the delivery of a 190 mg banana-flavored whole diet food pellet (Bioserv). Terminating key contact at other times did not produce food reward, and started a 6 s time-out condition during which the experiment was suspended and the monkey was required to refrain from contacting the response disk.

The response key was illuminated from behind, and this light was off during either the time-out condition or during a 5 s intertrial interval which followed reinforcement, and which allowed adequate time for the monkeys to eat. To signal the availability of the next trial, the key light pulsed at a rate of 5 pulses per s, when the individual securely contacted the response key, the light became lit continuously. The test stimuli were embedded within a broad band noise background that came on at the onset of each trial. The masking noise was presented at a level of 60 dBA.

Masked auditory thresholds were determined by the computer through the staircase or tracking method. In this procedure, the SPL of the stimulus to be presented in the up-coming trial is governed by the individual's response to the previous trial. If the previous trial was detected, the SPL of the stimulus for the next trial was reduced; if the monkey failed to detect the preceding trial, the converse occurred. Hence, the method tracks the minimal SPL of the stimulus required for the monkey to hear. Trained monkeys were initially presented with two different kinds of trials: test trails as described above (85%), and silent catch trials (15%) which permitted the measurement of the rate of guessing. A third type of trial presented clearly audible supraliminal stimuli; these trials always occurred immediately following the correct rejection of a catch trial, and they were programmed to ensure an adequate reinforcement density. Hence, the monkeys did not have to wait for very long intervals for the next detectable signal to be presented. This procedure has been found to produce a constant and uniform rate of responding.

Humans were tested with the same apparatus and with the same procedure as described above for the monkeys except that a wooden chair was substituted for the primate chair, and a click of the feeder served as feedback for a correct response, a food pellet was not delivered, and the intertrial intervals were reduced to 1 s in duration.

RESULTS

In Figure 1 we show the spectrum of the masking noise plotted in third octave bands. Though the acoustic source for the masker was white noise (equal energy per cycle), the acoustics of the room and speaker produced the resultant broad band spectrum. We compared this noise spectrum with the spectrum of rain forest, riverine forest and savanna habitats and found that the noise spectrum used here

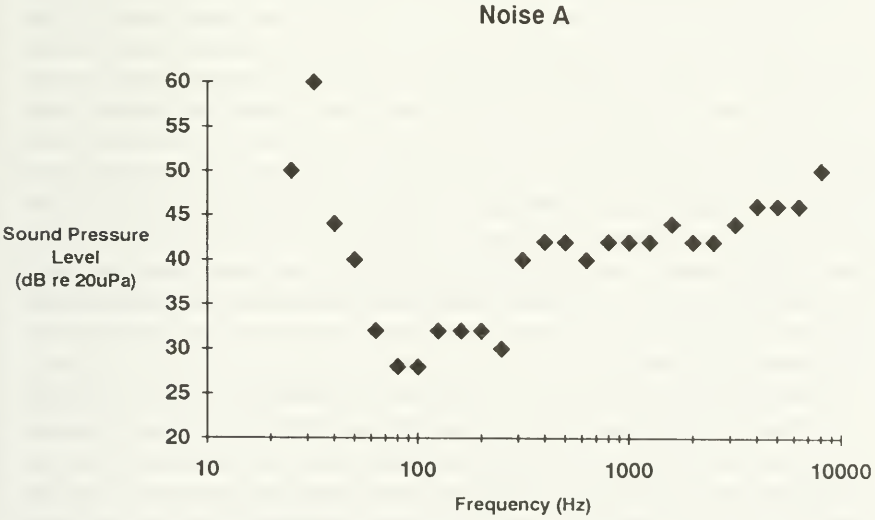

FIGURE 1. Spectrum in third-octave bands of the masking noise.

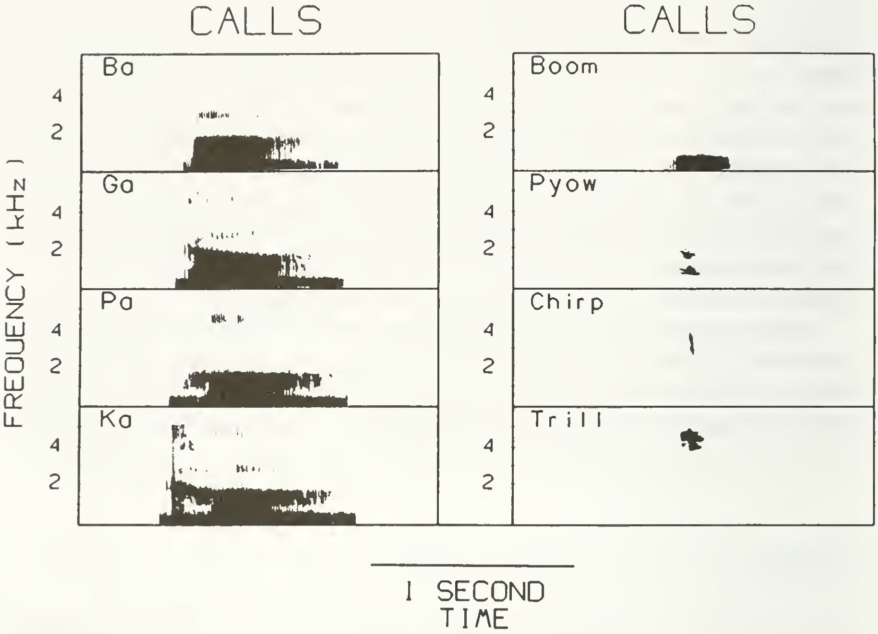

FIGURE 2. Sound spectrograms of the four human speech sounds (left panel) and the four blue monkey calls (right panel).

diverged strongly from those recorded in natural primate habitats (Waser & Brown, 1986).

Sound spectrograms of the eight test signals are presented in Figure 2. Signal duration ranged from 20 ms for the blue monkey *chirp*, to 880 ms for the *ka* speech utterance. The mean duration of the human speech sounds was 765 ms, while the mean duration for the monkey calls was only 115 ms. Hence, the human speech utterances were on the average 6.7 times longer than the blue monkey vocalizations. The sound spectrograms also reveal that the monkey vocalizations exhibited substantial diversity in dominant frequency. The *boom* call's dominant frequency was centered at 160 Hz, while the *chirp* and *trill* call were centered near 4 kHz. The dominant frequency band (Fo) of all four of the human speech sounds was centered near 1 kHz, and the speech sounds were much broader in bandwidth than the monkey vocalizations.

In Figure 3 we show the relative audibility of the test signals embedded in the broad band masking noise. The data shown here are averaged for six sessions. The performance of all listeners was very stable across sessions. The data for the humans are given by the open bars, and the data for the monkeys are displayed by the solid bars. There was very small variability among the listeners, the standard

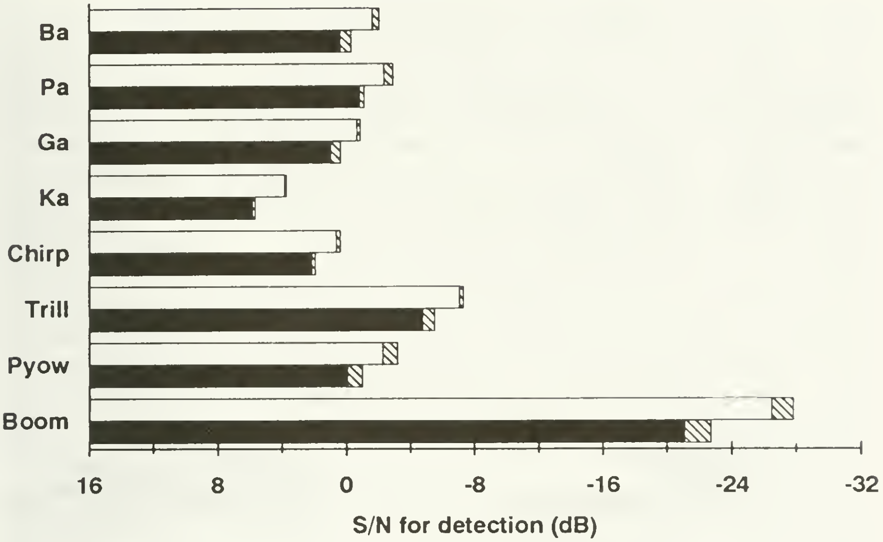

FIGURE 3. Signal-to-noise ratios for the eight test signals for human listeners (open bars) and monkey listeners (solid bars). Three individuals were in each group, and the standard error of the mean is denoted by the striped area at the end of each bar.

error of the mean was very small, and this is displayed by the striped area at the end of each bar.

The signal-to-noise ratio for detection ranged from 5.8 dB for blue monkeys presented with the *ka* sound to −26.5 dB for humans presented with the *boom* call. The results show that blue monkeys could detect these signals in noise nearly as well as human listeners. Humans were able to hear these signals on the average at a level that was 2.3 dB less than that heard by blue monkeys. The mean audibility of the speech sounds by humans was −0.25 dB, while the blue monkeys detected the same signals at a mean level of 1.53 dB. This yields a difference in sensitivity between humans and monkeys of 1.78 dB for the human speech sounds. The mean audibility of the monkey calls by humans was −8.85 dB, and the corresponding value for blue monkeys was −5.97 dB. This results in a difference in sensitivity between humans and blue monkeys of 2.88 dB for the blue monkey utterances. These findings are in keeping with other studies that show that humans are just a little bit more sensitive than nonhuman primates in most psychoacoustical tasks (Stebbins, 1973).

Pooling the data across species, the four monkey calls were audible at a mean signal-to-noise level that was 8.1 dB less than that required to hear the speech sounds. However, as is readily notable from inspection of Figure 3, most of this effect is due to the superior

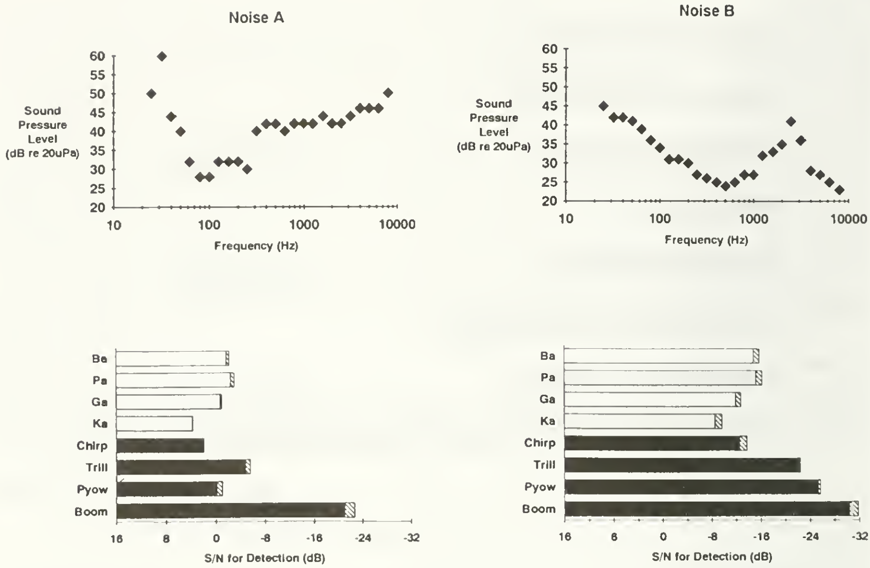

FIGURE 4. Signal-to-noise ratios for two noise spectra. The noise spectra are presented in the top panel. Noise A is the noise spectra presented here, Noise B is simulated rain forest masking noise (Brown, 1986). The corresponding signal-to-noise ratios for the eight test stimuli are given in the lower panel. Data are presented for conspecific listeners only.

audibility of the blue monkey *boom* vocalization. With the *boom* excluded from analysis, the mean difference in the audibility of the speech and monkey sounds was 2.6 dB. Hence, in the presence of the broadband masker used here, the blue monkey calls in general were slightly more detectable in noise than were the speech sounds, but this effect was not very strong.

DISCUSSION

These observations contrast with previous findings from our laboratory which used simulated rain forest noise as the masking sound (Brown, 1986). In Figure 4, top panel, we present spectra of the broadband masking noise used here, and the simulated rain forest masking noise used in earlier work. In the bottom panel of Figure 4 we show the relative audibility in noise of the 8 test stimuli detected by conspecific listeners in the two noise environments. Because Brown (1986) only tested the monkey calls with monkeys and the speech sounds with the humans, we plot the corresponding data in Figure 4.

Two observations are important here. First, the test signals were in general more audible when testing was conducted with the rain forest masker. Averaged across calls, the signal-to-noise ratio for the detection of this set of stimuli was -17.6 dB when testing was conducted with the rain forest masker, and -3.4 dB when testing was conducted with the broadband masker. Thus, the shape of the noise spectrum appears to have a strong effect on the relative audibility of complex signals in noise. Second, when testing was conducted with the rain forest masker, the monkey calls were audible at a relative level that was 10 dB less than that required for the detection of speech sounds (Miller, 1947; Hawkins & Stevens, 1950). As noted above, this effect was amplified by the superior audibility of the *boom* call. However, even after the *boom* was excluded from analysis, the monkey calls were still 7.4 dB more audible than the speech signals. This contrasts with the 2.6 dB value obtained above with the broadband masker. Thus, though the blue monkey calls examined here have a physical structure that enhances their audibility in noise relative to that for speech sounds, the results suggest that rather small variations in the spectral or amplitude characteristics of the noise present in the acoustic environment may have a strong influence in the audibility of vocal signals in nature.

Brown (1986) argued that forest monkeys had undergone selection to produce vocalizations that were more audible in noisy environments than were human speech sounds. However, it is unclear if the 10 dB difference reported by Brown (1986) was due to differences in the acoustics of monkey calls and human speech, the characteristics of the masking noise used, or differences in the abilities of humans and monkeys to detect complex sounds embedded in noise. The results of the present study show that blue monkeys are not in general better than human listeners in detecting complex signals in noise. Humans detected both monkey sounds and speech sounds at slightly lower signal-to-noise ratios than did the monkeys. It appears likely that humans would have performed similarly to monkeys in the rain forest masking noise (Noise B) if both species had been tested on both sets of stimuli. In contrast, the results suggest that differences in the audibility of speech and monkey sounds in the ecologically valid rain forest masking noise was due to both the spectral characteristics of the masking noise and species-specific differences in the structure of human and blue monkey vocal signals. Interestingly, the very brief blue monkey calls are more audible in noise for both human and monkey listeners than is the much longer duration human speech stimulus, and this effect is amplified when testing is conducted in noise that mimics the spectral characteristics of the natural environment.

As a rule, researchers in psychoacoustics have given little attention to the spectral and temporal characteristics of masking sounds used in perceptual studies (Brown, 1986; 1989a). Though researchers study the perception of biologically significant signals (e.g., vocalizations), they typically employ masking noises in these and in related studies which have no ecological validity or biological significance. Most researchers appear to tacitly assume that all masking sounds are approximately equivalent impediments to hearing. However, each natural habitat exhibits a constellation of acoustical characteristic that are likely unique unto itself (Brown & Waser, 1984; 1988; Waser & Brown 1986; Brown, 1989a; 1989b), and because communication systems evolved within a specific acoustical environment it is probable that the structure of some systems underwent selection to counter the impediments to acoustic communication characteristic of the local habitat. The observation that blue monkey calls are audible at lower signal-to-noise ratios when embedded in the rain forest masker compared to when embedded in a broad band masking noise that lacks ecological validity or biological significance is consistent with this hypothesis. The idea of a match between the acoustic structure of the vocal repertoire and habitat acoustics should be explored more fully by testing the audibility of species-specific vocal signals in primates resident in riverine and savanna habitats with ecologically appropriate and inappropriate masking noise.

How can it be that two different masking sounds with different spectral characteristics change the audibility of signals in noise? To our knowledge there has been no systematic study of the influence of various maskers on the audibility of signals in noise. However, several possible mechanisms merit comment. First, it is possible that the spectral density may be lower near the dominant frequency band of any test signal in one masking sound relative to that for another. For example, in Figure 4 the spectral density of Noise A is lowest at 80 Hz and 100Hz, while the spectral density of Noise B is lowest at 630Hz. These differences in the shape of the noise spectrum may account for the global differences in the audibility of the test signals embedded in the two noise samples. Many of the test signals have a dominant energy band near 1000 Hz, a region of low spectral density in Noise B, and it is possible that these signals are more audible in Noise B because they fall in a quiet zone within the spectrum of the noise. However, this explanation cannot account for the enhanced audibility of the *boom, chirp*, and *trill* vocalizations in Noise B. In these three cases, the spectral density of Noise B is approximately the same as that in Noise A at the dominant frequency region of the sound.

This observation suggests that differences in the audibility of various signals in different noise environments may not simply be due to a fortuitous congruence between “quiet zones” in the masker and the dominant frequency band of the signal. Brown (1989b) has suggested that the auditory system of many organisms may have been “designed” to promote the detection of signals in noise. Both psychoacoustical and physiological observations are pertinent to this hypothesis. For example, Young and Barta (1986) noted that the rate-intensity function of single units recorded in the cat were steeper if the test signals were presented in noise than if they were presented in the quiet. It was as if the auditory system functioned with greater precision under the biologically normal conditions of the presence of a background noise. Preliminary behavioral observations from our laboratory also reveal certain conditions in which the presence of noise improves perception. We have observed, contrary to our initial expectations, that monkeys reveal smaller frequency difference limens when tested at low sensation levels compared to larger sensation levels, or if tested at higher sensation levels when a masking sound is present. It is as if the monkey's auditory system "expects" the presence of masking noise, and it is "designed" to function best in the noise context.

These observations are consistent with the idea that the communication systems of some organisms have been selected to promote hearing in noise, and that the acoustical attributes of ecologically appropriate background noise may "interface" with the signal detection module of the auditory system in unexpected ways. In this perspective, it is as if each place in the natural world has its own special ambient sound, and that the ears of the organisms resident in each acoustic environment have undergone selection to be able to quickly discern the occurrence of a sound that does not belong, a sound which may signal the approach of a competitor, a predator, a relative, a mate, or one's prey. Classically the auditory system has been studied with biologically and ecologically irrelevant stimuli, pure tones, tone pips, clicks, noise bands, and the like. Though many investigators now study the auditory processing of species-specific communication sounds, few investigators are studying the perception of these signals in an ecologically "appropriate" acoustic environment. The data presented here suggest that a full understanding of the organization and design of the auditory system will require researchers to examine perceptual processes within the acoustic environment in which communication normally occurs and in which it evolved.

ACKNOWLEDGEMENTS

This work was supported by NIH grant R01 DC00164-11.

REFERENCES

Brenowitz, E.A. (1982). The active space of red-winged black-bird song. *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology, 147*, 511-522.

Brown, C. H. (1986). The perception of vocal signals by blue monkeys and grey-cheeked mangabeys. *Experimental Biology, 45*, 145-165.

Brown, C. H. (1989a). The active space of blue monkey and grey-cheeked mangabey vocalizations. *Animal Behaviour, 37*, 1023-1034.

Brown, C. H. (1989b). The acoustic ecology of East African primates and the perception of vocal signals by grey-cheeked mangabeys and blue monkeys. In R.J. Dooling & S.H. Hulse (Eds.), *The comparative psychology of audition: Perceiving complex sounds* (pp. 201-239). Hillsdale, NJ: Lawrence Earlbaum Associates, Inc.

Brown, C. H., & Schwagmeyer, P. L. (1984). The vocal range of alarm calls in thirteen-lined ground squirrels. *Zeitschrift für Tierpsychologie, 65*, 273-288.

Brown, C. H., & Waser, P. M. (1984). Hearing and communication in blue monkeys (*Cercopithecus mitis*). *Animal Behaviour, 32*, 66-75.

Brown, C. H., & Waser, P. M. (1988). Environmental influences on the structure of primate vocalizations. In D. Todt, P. Goedeking, & D. Symmes (Eds.), *Primate Vocal Communication* (pp. 51-66). Berlin: Springer-Verlag.

Byrne, R. W. (1981). Distant vocalizations of guinea baboons (*Papio papio*) in Senegal: An analysis of function. *Behaviour, 78*, 283-312.

Gautier, J-P., & Gautier, A. (1977). Communication in Old World monkeys. In T. Sebeok (Ed.), *How animals Communicate* (pp. 890-964). Bloomington, IN: Indiana University Press.

Hawkins, J. E., & Stevens, S. S. (1950). Masking of pure tones and speech by noise. *Journal of the Acoustical Society of America, 22*, 6-13.

Marler, P. (1973). A comparison of vocalizations of red-tailed monkeys and blue monkeys, *Cercopithecus ascanius* and *C. mitis* in Uganda. *Zeitschrift für Tierpsychologie, 33*, 223-247.

Miller, G. A. (1947). The masking of speech. *Psychological Bulletin, 44*, 105-129.

Morton, E. S. (1975). Ecological sources of selection on avian sounds. *American Naturalist, 109*, 17-34.

Moynihan, M. (1966). Communication in the titi monkey, *Callicebus*. *Journal of Zoology (London), 150*, 77-127.

Payne, R. S., & Webb, D. (1971). Orientation by means of long range acoustic signaling in baleen whales. *Annals of the New York Academy of Science, 188*, 110-141.

Ryan, M. J., & Brenowitz, E. A. (1985). The role of body size, phylogeny, and ambient noise in the evolution of bird song. *American Naturalist, 126*, 136-139.

Stebbins, W. C. (1973). Hearing of Old World monkeys (*Cercopithecinae*). *American Journal of Physical Anthropology, 38*, 357-364.

Waser, P. M., & Brown, C. H. (1986). Habitat acoustics and primate communication. *American Journal of Primatology, 10*, 135-154.

Young, E. D., & Barta, P. E. (1986). Rate responses of auditory nerve fibers to tones in noise near masked threshold. *Journal of the Acoustical Society of America, 79*, 426-442.