

Auditory looming perception in rhesus monkeys

Asif A. Ghazanfar*[†], John G. Neuhoff[‡], and Nikos K. Logothetis*

*Max Planck Institute for Biological Cybernetics, Spemanstrasse 38, 72076 Tübingen, Germany; and [‡]Department of Psychology, College of Wooster, Wooster, OH 44691

Edited by Peter Marler, University of California, Davis, CA, and approved October 1, 2002 (received for review August 6, 2002)

The detection of approaching objects can be crucial to the survival of an organism. The perception of looming has been studied extensively in the visual system, but remains largely unexplored in audition. Here we show a behavioral bias in rhesus monkeys orienting to “looming” sounds. As in humans, the bias occurred for harmonic tones (which can reliably indicate single sources), but not for broadband noise. These response biases to looming sounds are consistent with an evolved neural mechanism that processes approaching objects with priority.

Animals often use looming cues to respond rapidly to potentially dangerous approaching objects. Adaptive responses to both real and simulated visual looming have been shown in several taxa (1). For example, in both infant and adult rhesus monkeys (*Macaca mulatta*), a rapidly expanding circular shadow causes them to duck, flinch, jump back, or produce an alarm call, whereas a shrinking shadow causes no such fear or avoidance responses (2). It has been shown in some animal species that these rapid and adaptive responses to visual looming stimuli are the product of specialized neural circuits (e.g., refs. 3 and 4).

Vision plays an important role in notifying animals of danger. However, it is ineffective for detecting approaching objects that are out of sight. In addition to visual specializations, most animals have evolved parallel auditory warning systems to escape undesirable encounters. Indeed, the auditory system can provide information about hidden objects that could be imminently dangerous. It has been argued that a primary role of motion perception in the auditory systems is that of warning, either to direct the visual system toward the object, or to initiate appropriate avoidance behavior (5). In natural environments, many looming objects have an auditory component, and such approaching sound sources are characterized by dynamic increases in intensity (among other acoustic characteristics).

To date, only a few studies have investigated auditory looming perception in humans (6–11). However, a consistent finding across such studies has been a systematic underestimation of the time-to-arrival of the approaching source, and perceptual judgments of source distance indicating that the source is closer than actual. These results have been interpreted in terms of their evolutionary benefits in that they provide a “margin of safety” in the perception of looming objects. Rising intensity has been shown to be a particularly salient cue to source approach (9), and humans have been shown to reliably overestimate the change in level of rising-intensity tones compared with the estimates of level change in equivalent falling-intensity tones (7, 8). However, similar results are not typically obtained with broadband noise stimuli.

This spectral difference may reflect the reliability with which a sound can specify source identity and the ease with which a sound can be separated or parsed from background noise (8). Tones can specify source identity more reliably than noise, in part because of the correlation between the intensity changes of their individual spectral components. Listeners can use this correlation to parse auditory scenes and identify single sources as being separate from background noise (12). Of course, natural sound sources can produce both noise and harmonic tones. However, harmonic tones are more reliable markers for individual sources because correlated changes in harmonics are almost never produced by separately sounding sources. Con-

versely, broadband noise is often the result of multiple simultaneously sounding sources (13).

Together, these data suggest an adaptive bias for perceiving auditory looming that is associated with tonal intensity change produced by the approaching source. If such a bias in humans has in fact evolved because it provides a selective advantage, then it should also be present in a closely related species. However, auditory looming perception in the animal kingdom is relatively unexplored (but see refs. 14 and 15).

Our study addressed two fundamental questions: (i) Do nonhuman primates show a biased response toward rising-intensity (i.e., “looming” sounds) over falling-intensity sounds? and (ii) if so, is this bias spectrally dependent? We tested captive rhesus monkey subjects by measuring an adaptive and unambiguous head orientation response to unseen sound sources. We presented playbacks of rising- or falling-intensity sounds through a hidden speaker placed close to, and behind, the subject’s head, and measured the duration of the orienting response. We reasoned that, if rising intensity is a differentially salient environmental signal that can indicate a looming source, then orienting responses to such sounds should be longer in duration when compared with falling-intensity sounds (which can indicate a retreating sound source). If the direction of intensity change is not differentially salient, then subjects should show statistically identical orienting response durations.

Methods

Twenty male rhesus monkeys, part of a large colony of group-housed individuals at the Max Planck Institute for Biological Cybernetics, were used for this study. For each experiment, an individual was seated in a primate chair and brought to a small semianechoic room for testing. The subject sat with his back toward a black curtain that concealed a hidden speaker (a self-powered Advent AV570 speaker (Interact Accessories, Lake Mary, FL); frequency range: 40 Hz to 30 kHz \pm 3 dB). The speaker was positioned at head level, 75 cm away from, and to the right, of the subject. Each session consisted of two trials: a playback of a rising-intensity sound followed by a playback of falling-intensity sound (or vice versa; trial types were counterbalanced for order). Playbacks were controlled by using a laptop PC (Sony VAIO 505, Sony, Tokyo) using the SOUNDFORGE 4.5 software (Sonic Foundry, Madison, WI). Trials were initiated only if the subject had been looking forward for a minimum of 10 s.

In condition 1, 10 subjects were presented with a 1-kHz complex tone composed of a triangle waveform (Fig. 1A). In condition 2, another 10 subjects were presented with white noise (Fig. 2A). Both the tones and noise were 750 ms in duration and either rose exponentially in intensity from 65 to 85 dB or fell from 85 to 65 dB. They were sampled at 44.1 kHz, had 10-ms onset and offset ramps, and changed 20 dB in intensity from start to finish. The slopes of the changing sound amplitudes were exponential. Exponential changes more closely approximate the changes in intensity that occur because of constant source velocity in natural environments than do linear changes. Sound

This paper was submitted directly (Track II) to the PNAS office.

[†]To whom correspondence should be addressed. E-mail: asifg@tuebingen.mpg.de.

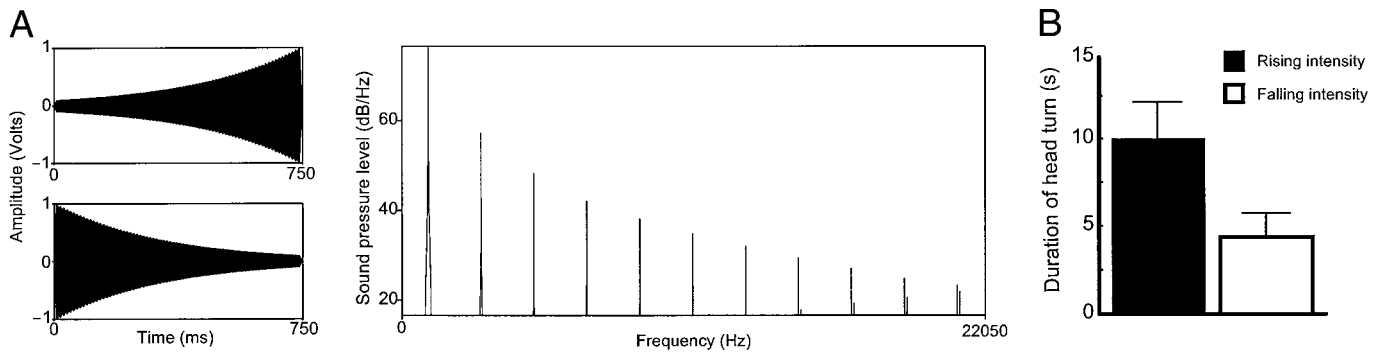


Fig. 1. Orienting responses to rising- and falling-intensity complex tone stimuli. (A) Time-amplitude waveforms and power spectrum of the complex tone (based on a 1-kHz triangle waveform) used in condition 1. (B) Duration of head-orienting responses to rising- and falling-intensity complex tones. Black bars represent responses to rising-intensity sounds, whereas white bars represent falling-intensity sounds. The y axis represents duration (in s) of first head turn as measured from the onset of the sound.

pressure levels were measured (at a distance of 75 cm) with a Brüel & Kjær 2238 Mediator sound level meter (Brüel & Kjær Instruments, Marlborough, MA) and a Brüel & Kjær 4188 condenser microphone.

Following playback of either trial type, subjects almost always oriented immediately toward the hidden speaker. All orienting responses to playbacks were videotaped by using a JVC (Tokyo) digital video camera (GR-DVL805) and digitally encoded onto a Dell Latitude C800 laptop computer (Round Rock, TX) by using the IEEE 1394a input and Adobe PREMIERE 6.0 software. The video acquisition rate was 30 frames per second with a frame size of 720×480 pixels. We measured response durations from the onset of the sound, to the time when the subject first began to turn his head back away from the speaker location. This response was unambiguous. Only the initial orienting response was measured, i.e., subsequent head turns were not included in the response measurement. All responses were scored blind to the trial type.

Results and Discussion

In condition 1, we tested 10 subjects by using a complex tone with a fundamental frequency of 1 kHz as our stimulus base. Subjects oriented to the rising-intensity tone significantly longer than to the falling-intensity tone. That is, despite the fact that duration, spectral content, and overall intensity were identical for the two sounds, subjects oriented on average for 9.99 ± 2.06 s (mean \pm SE) for the rising-intensity tone compared with 4.37 ± 1.39 s for the falling-intensity tone (two-tailed paired t test: $t(9) = 2.61$, $P = 0.028$; Fig. 1B). Furthermore, 10 of 10 subjects oriented longer after the rising- vs. falling-intensity tone (sign test, $P =$

0.001; Fig. 3A). These data suggest that rising-intensity tones are perceived as more salient than equivalent falling-intensity tones.

In condition 2, we tested whether the auditory looming bias depended on the harmonic characteristics of the stimulus. With a different set of 10 subjects, we used white noise instead of a complex tone, but with otherwise identical intensity parameters as those used in condition 1 (Fig. 2A). The results failed to show a significant difference between orienting responses to rising- and falling-intensity noise. For rising-intensity white noise, subjects' orienting duration was 4.56 ± 2.28 s, whereas for falling-intensity white noise, it was 3.86 ± 1.31 s [$t(9) = 0.280$, $P = 0.786$; Fig. 2B]. Three of 9 subjects oriented longer to the rising-intensity white noise (sign test, $P = 0.254$; one subject responded with equal duration for both trials) (Fig. 3B). These data demonstrate that the perceptual bias for rising-intensity sounds in rhesus monkeys (as in condition 1) depends on the spectral characteristics of the stimulus.

In a natural environment, an approaching sound source is characterized by a dynamic increase in intensity at the point of the listener. We have shown that rhesus monkeys show a strong orienting bias for a complex tone that increases in intensity compared with an equivalent tone that falls in intensity. This bias was not observed for white noise. Both the perceptual bias for rising intensity and its spectral dependency are strikingly consistent with results from human psychophysical studies (7, 8). The bias for rising intensity in both humans and rhesus monkeys could provide an increased margin of safety in perceiving looming sources. The pronounced effect for tonal sounds over noise may indicate adaptive priorities in the processing of auditory motion. Tonal harmonics that undergo correlated

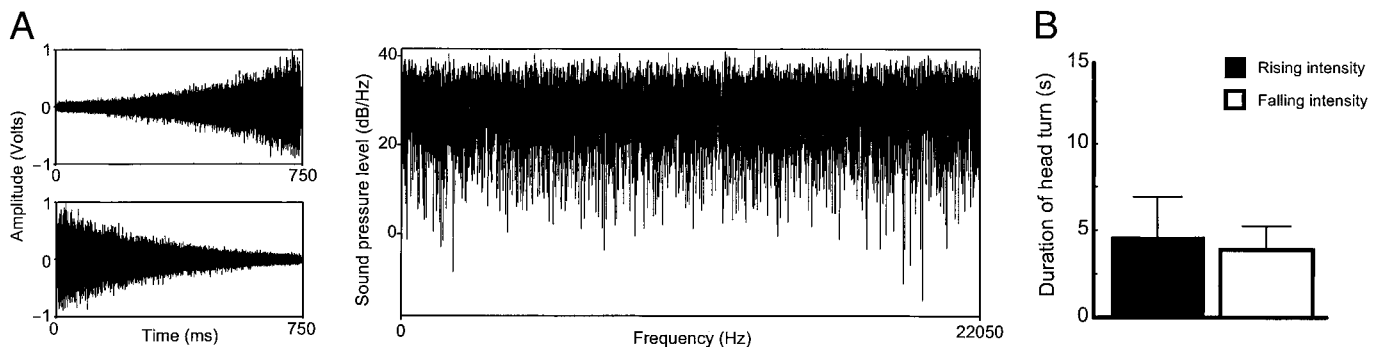


Fig. 2. Orienting responses to rising- and falling-intensity white noise stimuli. (A) Time-amplitude waveforms and power spectrum of the white noise stimuli used in condition 2. (B) Duration of head-orienting responses to rising- and falling-intensity white noise. Black bars represent responses to rising-intensity sounds, whereas white bars represent falling-intensity sounds. The y axis represents duration (in s) of first head turn as measured from the onset of the sound.

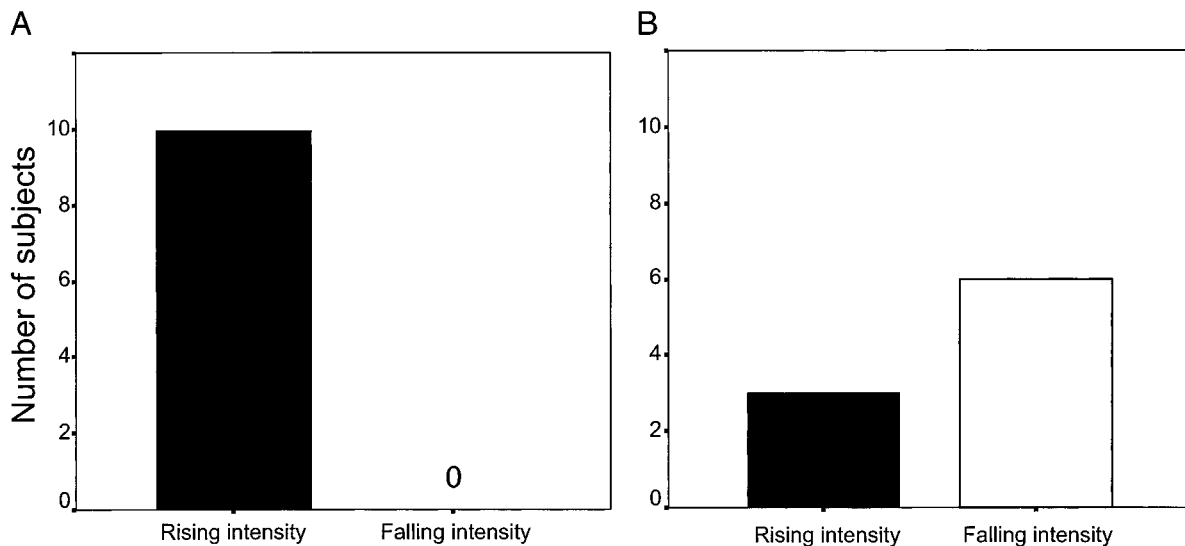


Fig. 3. Proportion of subjects responding longer to rising- vs. falling-intensity sounds. (A) Complex tone stimuli, condition 1. (B) White noise stimuli, condition 2.

changes in intensity can provide cues to source identity and location (12, 16, 17). The uncorrelated changes typical of the components of noise make parsing sources of noise more difficult (18, 19). When harmonic spectral cues are used to parse a source from an auditory scene, and intensity change is consistent with source motion, the processing of rising intensity appears to have priority over falling intensity.

The neural basis for this auditory perceptual anisotropy is currently unknown. However, a recent neurophysiological study in a primate model suggests that this perceptual bias may be cortical in origin (20). In marmoset monkeys, a greater proportion of primary auditory cortical neurons are selective for ramped (rising-intensity) sinusoids than damped (falling-intensity) sinusoids (20). However, the stimulus durations used in that study were much shorter than those used in the present study. We have also shown that the source spectrum can modulate responses to looming sounds. Single-cell recordings in the cochlear nucleus of cats, gerbils, and guinea pigs have shown that a complex interaction between units in the dorsal and

ventral auditory pathways is instrumental in the differential processing of tones and broadband noise at varying intensity levels (21–23). When such spectral differences are used to identify a sound source, and intensity change indicates source motion, the processing of rising intensity appears to take priority.

We conclude that rhesus monkeys, like humans, have an adaptive bias for perceiving looming, biologically salient sound sources. Previous work has suggested that the primary function of the auditory localization may not be to provide exact estimates of source location but, rather, to provide input to the listener's perceptual model of the environment (24). Thus, behavioral biases in response to auditory stimuli, as reported here, may be adaptive. The use of ecologically relevant, dynamic sound stimuli in auditory neuroscience research will likely reveal neural specializations for such behaviors (25).

We thank Marc Hauser, Cory Miller, and Laurie Santos for helpful comments and/or discussion. This work was supported by the Max Planck Society.

- Schiff, W. (1965) *Psychological Monographs: General and Applied* **79**, 1–26.
- Schiff, W., Caviness, J. A. & Gibson, J. J. (1962) *Science* **136**, 982–983.
- Hatsopolous, N., Gabbiani, F. & Laurent, G. (1995) *Science* **270**, 1000–1003.
- Sun, H. & Frost, B. J. (1998) *Nat. Neurosci.* **1**, 296–303.
- Guski, R. (1992) *Ecol. Psychol.* **4**, 189–197.
- Ashmead, D. H., Davis, D. L. & Northington, A. (1995) *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 239–256.
- Neuhoff, J. G. (1998) *Nature* **395**, 123–124.
- Neuhoff, J. G. (2001) *Ecol. Psychol.* **13**, 87–110.
- Rosenblum, L. D., Carello, C. & Pastore, R. E. (1987) *Perception* **162**, 175–186.
- Rosenblum, L. D., Wuestefeld, A. P. & Saldaña, H. M. (1993) *Perception* **22**, 1467–1482.
- Schiff, W. & Oldak, R. (1990) *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 303–316.
- Bregman, A. S. (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).
- Nelken, I., Rotman, Y. & Yosef, O. B. (1999) *Nature* **397**, 154–157.
- Lee, D. N., Simmons, J. A. & Saillant, P. A. (1995) *J. Comp. Physiol. A* **176**, 246–254.
- Lee, D. N., Vanderweel, F. R., Hitchcock, T., Matejowsky, E. & Pettigrew, J. D. (1992) *J. Comp. Physiol. A* **171**, 563–571.
- Ciocca, V., Bregman, A. S. & Capreol, K. L. (1992) *Q. J. Exp. Psychol. A* **44**, 577–593.
- Rogers, W. L. & Bregman, A. S. (1998) *Percept. Psychophys.* **60**, 1216–1227.
- Bregman, A. S., Abramson, J., Doehring, P. & Darwin, C. J. (1985) *Percept. Psychophys.* **37**, 483–493.
- Bregman, A. S., Levitan, R. & Liao, C. (1990) *Percept. Psychophys.* **47**, 68–73.
- Lu, T., Liang, L. & Wang, X. (2001) *J. Neurophysiol.* **85**, 2364–2380.
- Davis, K. A. & Voigt, H. F. (1997) *J. Neurophysiol.* **78**, 229–247.
- Nelken, I. & Young, E. D. (1994) *J. Neurophysiol.* **71**, 2446–2462.
- Palmer, A. R., Jiang, D. & Marshall, D. H. (1996) *J. Neurophysiol.* **75**, 780–794.
- Popper, A. N. & Fay, R. R. (1997) *Brain Behav. Evol.* **50**, 213–220.
- Ghazanfar, A. A. & Santos, L. R. (2002) in *Primate Audition: Ethology and Neurobiology*, ed. Ghazanfar, A. A. (CRC, Boca Raton, FL), pp. 1–12.