

Categorical Perception of Sound Frequency by Crickets

Robert A. Wytenbach, Michael L. May,* Ronald R. Hoy†

Partitioning continuously varying stimuli into categories is a fundamental problem of perception. One solution to this problem, categorical perception, is known primarily from human speech, but also occurs in other modalities and in some mammals and birds. Categorical perception was tested in crickets by using two paradigms of human psychophysics, labeling and habituation-dishabituation. The results show that crickets divide sound frequency categorically between attractive (<16 kilohertz) and repulsive (>16 kilohertz) sounds. There is sharp discrimination between these categories but no discrimination between different frequencies of ultrasound. This demonstration of categorical perception in an invertebrate suggests that categorical perception may be a basic and widespread feature of sensory systems, from humans to invertebrates.

Much of perception is continuous; that is, more stimuli can be discriminated than can be identified or labeled. For example, we easily discriminate colors differing minutely along the continuous dimension of wavelength even though we have a limited number of labels for ranges of wavelengths. Our labeling of color is categorical, in that there are sharp boundaries between our color labels, but our perception of it is continuous, in that we can discriminate small differences of wavelength within those boundaries (1).

Many sounds of human speech, on the other hand, are perceived categorically (2). For instance, the sounds /ba/ and /pa/ differ in one dimension, voice onset time (VOT). Although VOT can vary continuously, we label all sounds on that continuum as either /ba/ or /pa/ with an abrupt transition between them. Thus, our labeling, or identification, is categorical. What makes our perception of these sounds categorical is that not only do we give all of them one label or the other, we also perceive all of them as either /ba/ or /pa/. We do not normally perceive intermediates as sounding halfway between the two, despite a wide variation of VOT within each category (3).

Like us, Polynesian field crickets (*Teleogryllus oceanicus*) face the problem of responding differently to parts of a continuous spectrum of stimuli. In their case, this is sound frequency: Crickets call at 4 to 5 kHz, whereas bats produce ultrasound (typically 25 to 80 kHz) to echolocate their prey. A cricket in flight must make a critical choice between flying toward another cricket's call and escaping from a bat. Although the temporal pattern of sound pulses provides a cue

for the calling song, this choice can be made on the basis of frequency alone, because a temporal pattern that attracts at 4 to 5 kHz repels at ultrasonic frequencies (4).

A flying cricket's response to sound can be measured readily because of several directional body movements involved in steering (5). During escape turning, there is a fast initial swing of the hind leg into the path of the hind wing. During prolonged steering, either toward an attractive stimulus or away from a repulsive one, the forewings bank into the turn and the abdomen swings, rudderlike, toward the turn. By monitoring these movements, we were able to measure our subjects' responses to a variety of artificially generated sounds.

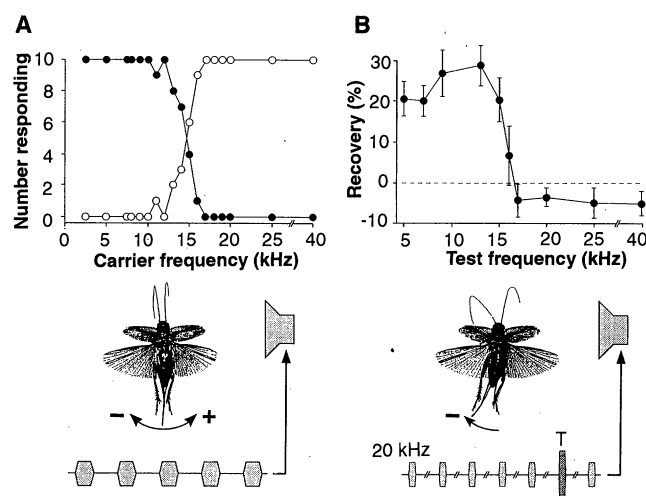
In 1970, Studdert-Kennedy *et al.* (6) proposed four operational criteria for absolute categorical perception of human speech: (i) distinct labeling categories with sharp boundaries, (ii) no discrimination between stimuli from the same category, (iii)

a peak in discrimination at category boundaries, and (iv) close agreement between actual discrimination performance and that predicted from the labeling results assuming absolute categorization.

We tested labeling in crickets by presenting synthesized calling-song-like patterns of sound pulses in which only the carrier frequency varied (7). If a cricket labeled a stimulus as attractive (in the calling-song category), it turned toward the speaker; if it labeled a stimulus as ultrasound, it turned away from the speaker. The third possibility, that intermediate frequencies would evoke no response, did not occur. The labeling function for 10 crickets (Fig. 1A) shows a sharp transition from attraction to escape between 13 and 16 kHz. Thus, the first criterion for categorical perception, distinct labeling categories, is met.

Testing categorical labeling across a continuous spectrum is relatively simple in animals and has been done several times (8), but determining the sensory discrimination ability of an animal is problematic. With adult humans, discrimination tests involve some variation of asking whether pairs of stimuli are the same or different and relying on verbal responses. Obviously, this cannot be done with nonverbal animals. Conditioning paradigms such as those used by Kuhl and Miller for chinchillas (9) and by Dooling and Brown for birds (10) come closest to this, with the subject motivated to do the best discrimination possible. With animals that cannot be trained, or when unconditional responses to natural stimuli are of interest, this becomes more difficult. Although some studies have attempted to infer discrimination from an animal's choice between two stimuli, this practice is

Fig. 1. Two tests of categorical perception. (A) Labeling. When presented with sound pulses in a calling-song-like temporal pattern, crickets turned toward (●) or away from (○) the speaker. Turns were of approximately the same magnitude at all frequencies. There is an abrupt shift from attraction to repulsion between 13 and 16 kHz, indicating a labeling boundary ($n = 10$; each cricket was tested once at each frequency). (B) Discrimination. The escape response was habituated by five pulses of 20 kHz (15). A subsequent pulse (T) at a test frequency caused the escape response to dishabituate only if it was below 16 kHz. This indicates that only frequencies below 16 kHz were discriminated from 20 kHz (mean $\pm$ SEM; $n = 5$ to 8 crickets per frequency). (Note the break in the carrier frequency scale between 25 and 40 kHz.)



Section of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853-2702, USA.

*Present address: American Scientist, Post Office Box 13975, Research Triangle Park, NC 27709, USA.

†To whom correspondence should be addressed. E-mail: rrh3@cornell.edu

questionable. The inference that two stimuli are discriminated if one is preferred to another may be valid, but the converse, that no preference means no discrimination, is not.

In research on human infants, habituation-dishabituation paradigms have been established to circumvent this problem (11); these paradigms have also been used to investigate discrimination in monkeys (12) and birds (13). The ultrasound-induced escape response of crickets decrements with repetition of the stimulus (habituation), and recovers (dishabituates) after presentation of a novel stimulus (14). Thus, dishabituation of the escape response can be used as a "same-different" test for crickets just as it has been used with other animals and human infants. A frequency that is discriminated from the habituating ultrasound will cause dishabituation, whereas a stimulus that is not discriminated from that frequency of ultrasound will not. This allows us to test frequency discrimination within the ultrasound category and between the ultrasound and song categories.

We presented five pulses of 20 kHz to habituate the escape response, after which we presented a pulse of a test frequency, followed by a final pulse of 20 kHz (15). The response to the fifth habituating pulse was reduced to 10 to 50% of the response to

the initial pulse (Fig. 2). The extent to which the response to the final pulse of 20 kHz dishabituated depended strongly on frequency of the test pulse (Fig. 1B). Only pulses below 16 kHz caused recovery when presented from the same location as the habituating series (16). Thus, there is sharp between-category discrimination but no discrimination within the ultrasound category. The dishabituation method cannot be used to test discrimination within the calling-song category, however, because attraction to the calling song builds up slowly and habituates irregularly if at all (4). The dishabituation test clearly demonstrates lack of discrimination within ultrasound but leaves open the question of whether there is continuous perception within the calling-song category.

Both tests of categorical perception show a categorical boundary between 13 and 16 kHz. This is an interesting and surprising result, given that this range is not of known behavioral significance to this species in nature (17). Communication calls of *Teleogryllus* are 4 to 5 kHz with relatively little energy at higher harmonics; echolocation cries of bats are between 25 and 80 kHz. There is a large gap between these two ranges, and one might expect either that the frequencies between them would elicit no behavioral response or that

attraction and avoidance would grade into each other over this range. It has been suggested that categorical perception, in general, functions to allow fast and accurate perceptual decisions (18); this is plausible in the case of the cricket, in which a rapid choice between attraction and escape is crucial to survival.

Categorical perception was initially thought to be unique to human speech (3, 19). However, categorical perception of nonspeech sounds (20) and visual stimuli (21) was soon found in humans. Subsequent studies showed that other mammals, including monkeys (12) and chinchillas (9), not only have categorical perception but that some of them respond categorically to attributes of human speech (22). More recently, evidence has been presented that swamp sparrows perceive certain calls categorically (13) and that budgerigars perceive some sounds of human speech categorically (10). There is also suggestive evidence from labeling tests that many other birds and frogs perceive their species-specific calls categorically (8). Our study suggests that there is room for further investigation of this phenomenon in the invertebrates as well as the lower vertebrates.

Two general conclusions may be drawn from this work. First, categorical perception may be a ubiquitous feature of perceptual systems in many animals, invertebrate as well as vertebrate, for simple as well as complex stimuli. Second, methods of human psychophysics have much to offer to the study of animal perception and cognition, in that there are well-established experimental paradigms that can be adapted to animal studies.

REFERENCES AND NOTES

1. R. M. Boynton, in *Handbook of Perception*, E. C. Carterette and M. P. Friedman, Eds. (Academic Press, New York, 1975) vol. 5, pp. 301-347.
2. A. M. Liberman, K. S. Harris, J. A. Kinney, H. Lane, *J. Exp. Psychol.* **61**, 379 (1961).
3. A. M. Liberman, F. S. Cooper, P. S. Shankweiler, M. Studdert-Kennedy, *Psychol. Rev.* **74**, 431 (1967).
4. A. Moiseff, G. S. Pollack, R. R. Hoy, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4052 (1978); T. G. Nolen and R. R. Hoy, *J. Comp. Physiol.* **159**, 423 (1986). In addition to being repelled by trains of ultrasonic pulses, flying crickets steer away from single pulses of ultrasound. Attraction to 4 to 5 kHz occurs only when the stimulus is presented as a train of pulses; a single pulse of this frequency elicits a behavioral response only at intensities much higher than those used here.
5. M. L. May, P. D. Brodfehner, R. R. Hoy, *J. Comp. Physiol.* **164**, 243 (1988); M. L. May and R. R. Hoy, *J. Exp. Biol.* **151**, 485 (1990).
6. M. Studdert-Kennedy, A. M. Liberman, K. S. Harris, F. S. Cooper, *Psychol. Rev.* **77**, 234 (1970).
7. G. S. Pollack and R. R. Hoy, *J. Comp. Physiol.* **144**, 367 (1981). The calling song of this species consists of a series of chirps and trills of 4.7 kHz, but these crickets respond equally well to a "continuous chirp" song. The song-like patterns we used in our experiments consisted of trains of 30-ms pulses with 5-ms rise and fall times at a rate of 15 pulses per second. All stimuli were presented at 80 dB SPL (sound pres-

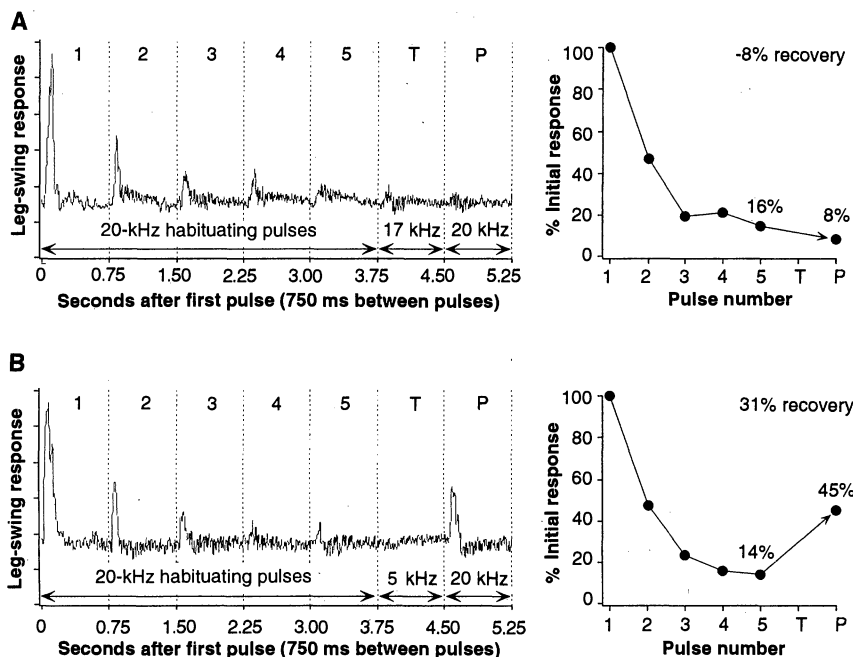


Fig. 2. Habituation and dishabituation. We presented five pulses of ultrasound (20 kHz) to habituate the escape response, followed by a test pulse of another frequency (T), followed by a final pulse of ultrasound (P). The magnitude of the leg-swing response was measured with a photocell (left panels; higher peaks indicate larger responses). All responses were then normalized to the first response (right panels). Recovery values shown in Fig. 1B were obtained by comparing normalized responses to the final pulse (P) and the fifth pulse. (A) When the test frequency is 17 kHz, the escape response continues to habituate, as indicated by the -8% recovery shown on the graph. (B) When the test frequency is 5 kHz, the escape response dishabituates, giving 31% recovery in this example.

- sure level) for 5 s. Tilt of the forewings was monitored to determine whether the cricket turned toward or away from the speaker. In these experiments, the amount of wing tilt did not vary with frequency. Over the range of intensities used here, the response to continuous-chirp stimuli varies with stimulus location rather than intensity [G. S. Pollack and N. Plourde, *J. Comp. Physiol.* **146**, 207 (1982)].
8. Reviewed by G. Ehret, in *Categorical Perception*, S. R. Harnad, Ed. (Cambridge Univ. Press, New York, 1987), chap. 10.
 9. P. K. Kuhl and J. D. Miller, *J. Acoust. Soc. Am.* **63**, 905 (1978).
 10. R. J. Dooling and S. D. Brown, *Percept. Psychophys.* **47**, 568 (1990).
 11. C. L. Miller and P. A. Morse, *J. Speech Hear. Res.* **19**, 578 (1976).
 12. P. A. Morse and C. T. Snowdon, *Percept. Psychophys.* **17**, 9 (1975).
 13. D. A. Nelson and P. Marler, *Science* **244**, 976 (1989).
 14. M. L. May and R. R. Hoy, *J. Exp. Biol.* **159**, 489 (1991).
 15. Stimuli consisted of single 10-ms pulses with 1-ms rise and fall times. A single pulse of 20 kHz was repeated five times with a 750-ms interval to let the escape response habituate. After another 750 ms, a single 10-ms pulse of the test frequency was presented, followed after another 750 ms by a final

- pulse of 20 kHz. Pulses of 20 kHz were 10 dB above the minimum required to elicit escape; the test pulse was another 10 dB greater than this amplitude. These amplitudes were chosen to ensure that all test stimuli, regardless of frequency, were clearly audible. The test pulse could come either from the same side as the habituating series or from the opposite side. The cricket's hind-leg steering response to each of the six 20-kHz pulses was measured with a photocell, as shown in the left panels of Fig. 2.
16. When the test pulse was replaced with silence, no recovery occurred. This shows that dishabituation at lower frequencies was not merely spontaneous recovery resulting from the absence of an ultrasonic pulse in the test pulse position. When presented from the side opposite the source of the five habituating pulses, a test pulse of any frequency caused dishabituation (of the same amount evoked by 5 to 15 kHz in Fig. 1B). This shows that all frequencies were audible and capable of producing dishabituation, despite differences in frequency sensitivity across the range of 5 to 40 kHz. The test pulses themselves, which were of equal amplitude caused steering responses of variable magnitude when they were >16 kHz. Responses were larger for 25- to 40-kHz test pulses than for 17- to 20-kHz test pulses. The difference is due to the greater sensitivity (~10 dB) of the escape response to 25 to 40

- kHz than to 17 to 20 kHz (4) and does not imply that the frequencies were discriminated.
17. F. Libersat, J. A. Murray, R. R. Hoy, *J. Comp. Physiol.* **174**, 485 (1994).
 18. E. Rosch, in *Cognition and Categorization*, E. Rosch and B. B. Lloyd, Eds. (Erlbaum, Hillsdale, NJ, 1978), pp. 27–48; R. J. Herrnstein, in *Animal Cognition*, H. L. Roitblat, T. G. Bever, H. S. Terrace, Eds. (Erlbaum, Hillsdale, NJ, 1984), pp. 233–261.
 19. I. G. Mattingly, *Am. Sci.* **60**, 327 (1972).
 20. J. D. Miller, C. C. Wier, R. E. Pastore, W. J. Kelly, R. J. Dooling, *J. Acoust. Soc. Am.* **60**, 410 (1976).
 21. R. E. Pastore et al., *J. Exp. Psychol. Hum. Percept. Perform.* **3**, 686 (1977).
 22. P. K. Kuhl, *J. Acoust. Soc. Am.* **70**, 340 (1981).
 23. We especially thank E. Spelke for suggesting the dishabituation paradigm to R.R.H. We also thank E. Spelke, J. Cutting, and members of the Hoy lab for helpful comments on the manuscript. Supported by National Institute of Deafness and Communicative Disorders grant R01-CD00103 and National Institute of Mental Health (NIMH) grant K05-MH1148 to R.R.H., NSF graduate fellowship and NIMH graduate training grant T32-MN15793 to R.A.W., and NIH graduate training grant T32-GM07469 to M.L.M.

18 April 1996; accepted 12 July 1996

Mechanisms of Heading Perception in Primate Visual Cortex

David C. Bradley, Marsha Maxwell, Richard A. Andersen,*
Martin S. Banks, Krishna V. Shenoy

When we move forward while walking or driving, what we see appears to expand. The center or focus of this expansion tells us our direction of self-motion, or heading, as long as our eyes are still. However, if our eyes move, as when tracking a nearby object on the ground, the retinal image is disrupted and the focus is shifted away from the heading. Neurons in primate dorso-medial superior temporal area responded selectively to an expansion focus in a certain part of the visual field, and this selective region shifted during tracking eye movements in a way that compensated for the retinal focus shift. Therefore, these neurons account for the effect of eye movements on what we see as we travel forward through the world.

When we move along a straight path, the retinal image appears to expand, creating a vector field on the retina in which all vector directions point away from the focus of expansion. When the eyes are still, this focus corresponds to the direction of self-motion (the heading), which humans can easily identify (1, 2). When the eyes move, the problem becomes more complicated. If we turn our eyes leftward to maintain fixation on an object to the left of our motion path, a rightward laminar flow component is added to the expanding retinal field. This combined field is similar to the original expansion field (when the scene contains little depth variation), but the apparent

focus position becomes shifted to the left of the heading (Fig. 1). Humans can still estimate heading during pursuit eye movements (2, 3), showing that they can correct for this shift. Recent psychophysical experiments have shown that this correction requires eye-movement information (3, 4). One might therefore expect the brain site or sites responsible for heading computation to process smooth-pursuit eye-movement signals as well as optic flow.

We examined the dorso-medial superior temporal (MSTd) area because its neurons are responsive to optic-flow stimuli (for example, expansion) (5–7) as well as to smooth-pursuit eye movements (7–9). Many MSTd cells are also selective for the position of the focus of expansion (10). Our goal was to determine how optic flow and eye velocity signals might be used in area MSTd to compute the direction of self-motion.

We trained a rhesus monkey to fixate on a point on a computer screen (11). The point was either stationary or was moved at a constant velocity of 15.7°/s (12). While the monkey fixated, we showed a large, expanding random-dot pattern for 1 s. The position of the focus in this expansion was varied along an axis to simulate different directions of self-motion, or headings. This axis was made parallel with each neuron's preferred pursuit direction (that eliciting the strongest response), which was determined in preliminary tests (13).

During the stimulus presentation we recorded single-neuron activity with a micro-electrode inserted in MSTd cortex (14). The optic flow preference (that is, expansion, contraction, or rotation) of each MSTd cell was first determined. If the cell responded best to expansion, we carried out the following tests (cells preferring contraction and rotation were also studied and are discussed below). First, while the monkey fixated a stationary point, we recorded the cell's firing rate as a function of the position of the flow focus on the screen (the simulated heading) (15). This allowed us to measure the neu-

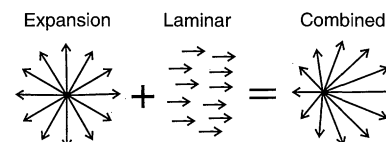


Fig. 1. When we move forward, the visual world appears to expand. If the eyes are still, the focus position tells us our direction of heading. However, leftward eye movement adds rightward laminar flow motion to the retinal image, which shifts the focus. To recover the heading direction, we must correct for this focus shift.

D. C. Bradley, M. Maxwell, R. A. Andersen, K. V. Shenoy, Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.
M. S. Banks, School of Optometry, University of California, Berkeley, CA 94720–2020, USA.

*To whom correspondence should be addressed.