THREE SONARS IN ONE BAT: LAYERED BIOSONAR IMAGES WITH UNIFIED SPATIAL DIMENSIONS

JA Simmons Dept. of Neuroscience, Brown University, Providence, RI 02912 USA, and Bio-navigation Research Center, Doshisha University, Kyotanabe 610-0321, Japan
MA Bates Dept. of Psychology, Brown University, Providence, RI 02912 USA
S Hiryu Faculty of Life and Medical Sciences, Doshisha University, and Bio-navigation Research Center, Doshisha University, Kyotanabe 610-0321, Japan
H Riquimaroux Faculty of Life and Medical Sciences, Doshisha University, and Bio-navigation Research Center, Doshisha University, Kyotanabe 610-0321, Japan

1 INTRODUCTION

The big brown bat (*Eptesicus fuscus*) is an insectivorous species that uses echolocation to perceive objects in its surroundings. Big brown bats transmit frequency-modulated (FM) sounds containing several harmonics that, collectively, span frequencies from about 20 to 110 kHz. Echolocation sounds have very high sound pressures, about 100-130 dB SPL (p-to-p) close to the open mouth. Broadcast durations are short and are adjusted to approximately match the two-way travel time of echoes from the target of immediate interest so that echoes do not overlap broadcasts. Durations vary from about 10-20 ms when the bat is flying in the open, to 2-5 ms when the bat is approaching within about 1meter of objects, and to about 0.5 ms in the terminal part of an interception. Interpulse intervals are scaled to the larger scope of the environment, ordinarily extending to at least 50-100 ms so that echoes from all objects in the scene have returned before the next sound is sent out, with the caveat that sounds often are grouped into pairs or triplets, called strobe groups" created by emitting sounds at alternating short and long intervals, which presumably serves to organize the bat's perception of auditory target "scenes." Figure 1 illustrates spectrograms for a series of biosonar sounds recorded during the late approach and terminal stages of target interception by a big brown bat. The 1st harmonic (FM1) sweeps from about 55 kHz down to 23 kHz, and the 2nd harmonic (FM2) sweeps from about 80-90 kHz down to 50 kHz (Figure 1). A short segment from the 3rd harmonic (FM3) usually is present, too, sweeping down from about 100 to 80 kHz. The harmonics give the spectrograms in Figure 1 a fingerprint-like appearance that bats use for recognition of their echoes.

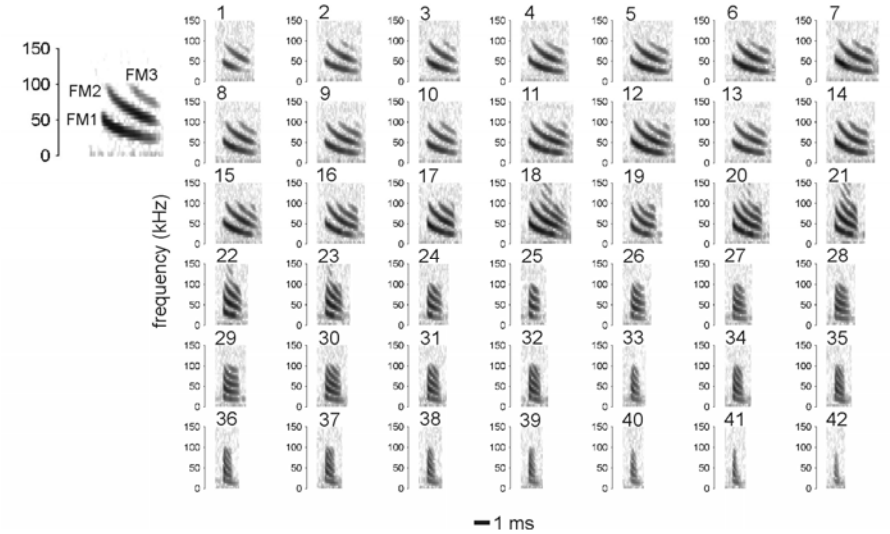

Figure 1: Spectrograms of biosonar sounds

FM biosonar sounds of big brown bats have several harmonics. These spectrograms show a sequence of echolocation sounds recorded from a flying big brown bat during interception of a microphone that served as a convenient target to obtain high fidelity recordings of the full broadcast bandwidth [3]. All of these sounds contain frequencies up to 100-110 kHz, with 3-6 harmonics (*e.g.*, FM1, FM2, FM3). The spectrograms have a fingerprint-like quality that makes each sound distinctive from neighboring sounds in the sequence. In this paper we examine how the bat takes advantage of small differences in the spectrograms of echoes to distinguish echoes of its own broadcasts from the sounds or echoes of other bats, and distinguish echoes reflected by the target of immediate interest from echoes reflected by objects in the background, or clutter. The trail of explanation is somewhat complicated because it involves the intertwining of the bat's perceptual dimensions with auditory mechanisms of echolocation and with target acoustics.

2 THE AUDITORY RECEIVER

The concept of "display" is useful for thinking about how the bat's sonar receiver is organized. In ordinary parlance, a display presents the human operator with images of target locations derived from the processing of echoes. Two long-familiar examples are the range-axis oscilloscope trace ("A-scope" display of echo delay by echo strength) and the polar plot of target range and horizontal direction (plan position indicator, or "PPI" display). These displays each have two spatial dimensions, plus the evolution of their spatial dimensions over time, from one broadcast to the next, while the target and the sonar move relative to each other. Sometimes the progression of new images over time is itself displayed spatially, as in side-scan sonar, where successive A-scope traces are plotted adjacent to each other to create a 3-dimensional mosaic image of range and cross-range dimensions plus echo strength. In echolocation, the bat's auditory system creates displays, too, but the spatial and temporal dimensions both of echoes and of images are folded together in an unconventional way. For the basic display, target range is transformed from echo delay by neuronal delay-lines and coincidence detectors into a representation based on "delay-tuned" neurons that seems crudely analogous to an A-scope display. Different neurons encode different ranges, but single-cell precision is crude, and there is considerable overlap between neurons separated by delay differences smaller than the size of their tuned delays themselves. Moreover, in big brown bats, this display is not static—it does not persist in the same neurons throughout the interval between one broadcast and the next. Instead, the A-scope-like display of range shifts around over time to appear in different groups of neurons every few hundred microseconds. The dynamic nature of this delay representation suggests that the progression of the display over time from one group of neurons to a succession of other groups might be exploited to continue processing the image after the echo has been received rather than simply to produce a stationary image for each broadcast that persists in time. It might, for example, facilitate processing of echoes that arrive at different delays from multiple targets because newly-arriving echoes are not competing momentarily for the same display elements as would be echoes that already have been received, and whose representation has since passed onward to different groups of neurons. This is but one example of the loose nature of the analogy between conventional sonar displays and the auditory system's own representations. Auditory and sonar displays seem to serve similar functions, such as determining where targets are localized and what features are perceived, but it is not so easy to make a more complete comparison, not least due to the elusive nature of the display-screen's "observer" in the bat's case.

2.1 AUDITORY REGISTRATION OF BROADCASTS AND ECHOES

The bat's auditory system serves as the sonar receiver for processing the information conveyed in echoes and transforming it into perceived images that depict the locations and features of objects. These perceptions arise from a cascade of processing stages that begin when the inner ear registers the time-of-occurrence of successive frequencies in the FM sweeps of broadcasts or echoes. Neurons in the auditory nerve receive their excitation from auditory receptor cells in the cochlea; they constitute the first stage in this cascade. In big brown bats, these first-stage neurons are tuned to frequencies from about 10 kHz to 100 kHz. For each neuron, the specific frequency depends on the location of that neuron's receptor cell along the frequency map of the spiral cochlea. Individual neurons at virtually all subsequent (*i.e.*, second and higher) stages of processing are tuned to specific frequencies in the 10-100 kHz band, too. They derive their tuning from excitatory inputs that originate in the tuning of the first-stage neurons of the auditory nerve. This excitation can be traced from stage-to-stage through progressively higher levels of processing. When the bat's ears are stimulated with short-duration FM sounds that mimic short-duration (1-5 ms) FM broadcasts (see Figure 1) and echoes, neurons throughout the bat's auditory pathways mostly respond to the occurrence of their tuned frequencies in the FM sweeps with single spikes for each sound. If one considers the various frequency-tuned cells at each processing stage as an array of frequency sampling points, then the volley of single spikes occurring across these neurons registers the shapes of the FM sweeps by their timing, or latency. First, each biosonar broadcast is received at the bat's ears when it is sent out, and a volley of spikes is initiated at the cochlea, to travel from stage to stage upward along auditory neuronal pathways. Then, echoes of that broadcast return to the bat's ears at different delays according to the distance to the objects that reflect them. Bats perceive the distance to targets from these delays. The new volley of spikes evoked by each of the incoming echoes also travels upward along the auditory pathways, "in pursuit of" the volley of responses previously evoked by the broadcast. The auditory representation of the target scene as a whole (broadcast plus its echoes) thus consists of a series of volleys of single spikes, each of which traces the shape of the FM sweeps across both frequency (from the tuning of the neurons) and time (from spike latency). In that these volleys have dimensions of frequency and time, they really are spectrograms—almost. Each spike that occurs in any of the neurons in the processing chain originates from a spike triggered by a threshold-crossing event at the first stage of neuronal coding, where excitation produced by auditory receptor cells tuned to different frequencies evokes spikes in their corresponding first-stage neurons. As with any threshold-triggering process, changes in the amplitude of the analog signals lead to changes in threshold-crossing time. For big brown bats, the latencies of spikes are retarded (by about 15 μs/dB) if the sound is attenuated and advanced (by about 15 μs/dB) if the sound is increased in strength. Due to this "amplitude-latency trading," big brown bats perceive changes in echo amplitude applied uniformly across the echo spectrum as corresponding changes in echo delay, or target range, at a trading ratio of -15 μs/dB. If amplitude changes are not applied uniformly to all echo frequencies but instead to some echo frequencies and not others, the neuronal spike patterns that make up the bat's FM auditory spectrograms are warped along their time axis. This warping occurs because the spectral level of the echoes at some frequencies does not exactly match the corresponding spectral levels in the broadcast, causing spikes registering the times-of-occurrence of different frequencies undergo different amounts of amplitude-latency trading. For example, neuronal spectrograms for echoes that arrive at the bat's ears after being lowpass filtered relative to the broadcast (by atmospheric absorption or due to off-axis location of the target) are slightly stretched in time by the progressively greater spike latencies caused by lower amplitudes at higher frequencies. Such frequency-dependent warping affects the bat's perception of the delay of echoes because the bat uses the spike patterns to extract echo delay. If we view spectrogram recognition as the process underlying the bat's determination of which sounds actually are echoes, it becomes critical to learn how sensitive bats are to distortions of the shape of the neural spectrograms of echoes in relation to the spectrogram of the preceding broadcast.

2.2 AUDITORY DISPLAYS

Higher-level stages of neural representation in the bat's auditory processing cascade can be thought of as displays for bringing together information to control various kinds of responses appropriate for each stage. The first such display conditions the neuronal spectrogram representation for subsequent use as well as controls low-level reflexes to coordinate transmission and reception of biosonar signals. This first functional stage of the neural processing cascade is located in the auditory brainstem. It regulates the occurrence of such rapidly-acting reflexes as contraction of the middle-ear muscles when intense sounds are received. These same muscles also contract in synchrony with the emission of sonar broadcasts to act as a gain control for echoes arriving at different delays. Laryngeal mechanisms for transmitting the broadcasts and auditory mechanisms for regulating the gain of the signals reaching the inner ear thus are linked together by receiving correlated neuronal commands. The auditory brainstem also processes the neuronal spectrograms arriving along auditory-nerve fibers by sharpening the precision of the time and the frequency dimensions for use by higher stages, as will be described below.

The second functional auditory display is located in the midbrain inferior colliculus. There, subpopulations of neurons tuned to different individual frequencies in the 10-100 kHz band produce single-spike responses along a continuum of different latencies rather than immediately upon arrival of their inputs from the brainstem. In the inferior colliculus, the span of latencies ranges from about 4 ms, when the arrival of inputs to the inferior colliculus from the auditory brainstem causes the first spikes to occur at each frequency, to progressively longer latencies—up to 10-15 ms at the higher frequencies of FM2 and up to 30-50 ms at the lower frequencies of FM1. Systematic retardation of response latency across different cells tuned to the same frequency is achieved by applying excitatory inputs only to those cells with short latencies (~4-6 ms). The majority of cells receive instead inhibitory inputs from the brainstem whose strength and duration are scaled to prevent a spike from occurring for a specific silent interval (most from ~5-35 ms) that is different in each cell. Different cells in the inferior colliculus thus produce single-spike responses at their own particular characteristic latencies only after their inhibitory inputs have waned. Consequently, the volley of responses representing the FM sweeps in a bat-like sound shifts from cell to cell as the inhibitory periods decay in one cell after another until the entire latency sequence is finished. This means that, at different moments, different spikes and different cells actually embody the volley representing the sound. In effect, the location of the spikes comprising each volley moves from cell to cell as time evolves. At any given moment, across cells tuned to different frequencies, the pattern of spike latencies retains a kind of trace or template for the shape of the FM sweeps in the original sound. For each outgoing sonar broadcast, this neuronal trace persists in the inferior colliculus for about 10-30 ms and thus is available to be compared with similar neuronal traces representing the FM sweeps of incoming echoes when their responses make their way up to the inferior colliculus, too. As time passes after transmission of the broadcast, and echoes start to arrive at the bat's ears, the whole scene in front of the bat comes to be displayed by the volleys of spikes moving across subpopulations of neurons in the inferior colliculus tuned to different frequencies. For a complex scene composed of multiple objects at different distances, the volley of spikes representing the spectrogram of the broadcast shifts from short-latency neurons to progressively longer-latency neurons until the longest-latency cells have responded and the display comes to an end. Similar volleys of spikes representing the spectrograms of echoes follow the volley for the broadcast across the latency sequences in the inferior colliculus. Because individual spikes in all of these volleys are traveling from one cell to the next in a determined sequence, the subpopulation of neurons at each frequency takes on the appearance of a delay-line. However, propagation of these spikes across the inferior colliculus actually is regulated not by connections from cell-to-cell in series but by different strengths and durations of the inhibitory inputs delivered to the inferior colliculus more or less simultaneously from the auditory brainstem. That is, although excitation does not actually travel from one cell of the inferior colliculus to the next, as if each frequency-tuned subpopulation were a transmission line with different delay taps, the mechanism as a whole nevertheless is analogous to sets of delay-lines at individual frequencies because the responses "propagate" regularly across the subpopulation of cells tuned to each frequency. Hypothetically, if responses passing through the parallel neuronal subpopulations in the inferior colliculus could be frozen in time at any given moment, the spectrograms of the broadcast and the echoes, as well as the time that has elapsed between the broadcast and each echo, could be "viewed," provided one knows the tuned frequencies and the characteristic latencies of individual cells. It is in this sense that we treat the neuronal representation in the inferior colliculus as a display of target range.

The simultaneous presence of ongoing responses to the broadcast and to echoes of that broadcast in the inferior colliculus sets the stage for the third auditory display, which expresses its images at the bat's auditory cortex based on information extracted from the display at the inferior colliculus.

The principle that underlies the “read out” of echo delays from the inferior colliculus is illustrated by the hypothetical freezing of responses mentioned above. At any particular frozen moment, responses occurring simultaneously both to the broadcast and to each of its echoes are captured in time; by freezing the responses at successive moments, the history of the volleys of responses can be followed as it shifts from cell-to-cell within the inferior colliculus. The simultaneity of these responses is detected by neurons that register coincidences between spikes evoked by the broadcast, and stored in the delay-lines of the inferior colliculus, and spikes evoked by incoming echoes. If the responses progressing through time in the inferior colliculus (that is, moving along the “delay-lines”) are considered as a running display of spectrograms for each broadcast and its associated series of echoes, then the responses in the cortex can be considered as “reading out” that display by keeping track of which neurons produce responses to echoes at the same moment that other neurons are producing responses to the broadcast. Neurons in the auditory cortex respond only if echoes are present at the delays to which they are tuned by virtue of which combinations of neurons from the inferior colliculus supply their inputs. There actually is a parallel third display in the superior colliculus, a place that in most mammals controls a variety of tracking responses such as following objects with the orientation of the eyes and head. In big brown bats, the superior colliculus receives echo-delay information similar to that supplied to the auditory cortex, and it exerts control over the orientation of the bat’s head and ears, while also regulating the duration and other features of biosonar broadcasts that track the locations of targets. As the cortex is usually described as initiating voluntary movements, whereas the superior colliculus controls a suite of orienting and tracking reflexes, it is widely assumed, albeit without much truly direct evidence, that what the bat “perceives” (the images developed by its sonar system) is determined by neuronal activity taking place in the auditory cortex. Given that this is a sonar system, one function of the highest level of display obviously must be to portray the locations and identifying features of any targets present in the vicinity—the target scene, but it also can be expected to participate in other functions such as distinguishing between a target of particular interest and other objects in the background—clutter suppression.

2.3 TARGET ACOUSTICS AND DISPLAY OF RANGE AND SHAPE

Observations of bats flying in vegetation while chasing each other or searching for prey shows that suppression of clutter interference must be an important part of echolocation. To understand how bats approach the problem of clutter, a different aspect of the third, cortical display of information carried by biosonar echoes has to be considered. The place to start is the acoustics of target echoes (Figure 2).The simplest type of target is a point-reflector located straight ahead of the bat and not farther away than a meter or so. The sound that impinges on such an “ideal” target is reflected with approximately equal strength at frequencies from 20 to 100 kHz, so that its echo is a just delayed replica of the broadcast. However, the frequency content of the incident sound itself is affected by the distance traveled out to the target and back. For biosonar sounds produced by big brown bats, atmospheric absorption is greater at the higher frequencies of FM2 and FM3 than the lower frequencies of FM1. Consequently, the farther the sound has to travel, the greater is the lowpass filtering it experiences on its journey (Figure 2A). For example, echoes returning from a target located 5 m away will have FM2 attenuated by about 6-7 dB relative to FM1. Moreover, the location of the target in the broadcast beam has an analogous lowpass effect on echoes. Because the transmitted beam is very wide at low frequencies but becomes gradually narrower at higher frequencies, only a target located straight ahead of the bat’s open mouth will be ensonified with the full bandwidth of the broadcast. For targets located off-axis, the incident sound will lack its higher frequencies (FM2, FM3) and contain only the lower frequencies (FM1) with any degree of fidelity (Figure 2B). By tracking the target of interest with the aim of its head, the bat keeps the target on the center of the broadcast beam, thus insuring that it will be ensonified with the full transmitted bandwidth. Unless the target is located in front of the bat’s head and not too far away, its echoes will always appear to be lowpass versions of the broadcast. In contrast, to the global effect of lowpass filtering, echoes from a point-target being tracked at ranges nearer than 1-2 m will contain all of the frequencies present in the broadcast. Insects, however, are not ideal point reflectors but instead typically consist of a small number of prominent reflecting parts (“glints”)—head, wings, abdomen—that largely determine the spectrum of echoes. The target’s particular contribution to echo structure is due to its range, for echo delay, and to its shape, for determining the spectrum through interference between reflections from its component glints. The example of insect echoes in Figure 2C shows a waterfall display of spectra from a succession of 126 echoes reflected by a fluttering moth. The waterfall display depicts the wingbeats themselves as periodic bursts of high-amplitude returns at the higher frequencies of FM2 and FM3 whenever the insect's wings are momentarily aligned perpendicular to the incident sound, while FM1 remains relatively stable in strength across wingbeats. Aside from the broad effects of the wings "flashing" in the sound beam as they beat, each individual spectral slice in Figure 2C contains a series of peaks and notches due to overlap and interference between reflections from the moving glints. In contrast to the highly variable individual spectra, the average spectrum collected across all of the echoes (Figure 2D) is only mildly lowpass filtered because the variable delays between the glint reflections have a somewhat stronger cancellation effect at the shorter wavelengths of FM2 and FM3; FM1 is less influenced by the interference because the longer wavelengths are closer to the dimensions of the target itself. From one echo to the next, the notches shift back and forth across different frequencies according to the separation of the glints along the range axis, or to the separation of their reflections in delay. These notches are registered by the bat's auditory system and used to estimate the delay separation of the reflections, and, thus, the spatial arrangement of the glints. In the third sonar display, at the bat's auditory cortex, neurons not only respond selectively to echo delay, for target range, but they respond selectively for the presence of notches in the echo spectrum at the particular frequency to which they are tuned. The cortical display thus expresses the orthogonal dimensions of overall echo delay in the time domain and glint spacing in the frequency domain to create a combined representation of range and shape.

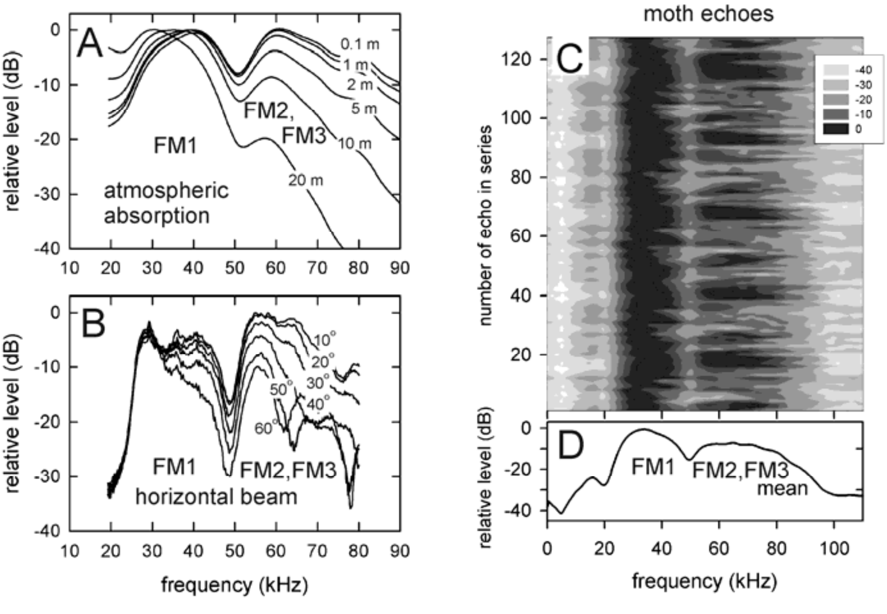

Figure 2: Acoustics of targets in air

(A) Relative spectra for FM echoes from an ideal target located at different distances from a big brown bat. The progressive lowpass filtering caused by propagation over greater distances imposes differential atmospheric attenuation on FM2 and FM3 compared to FM1. (B) The bat's broadcast beam is wider at frequencies of FM1 than FM2 or FM3, leading to progressive loss of FM2 and FM3 relative to FM1 for greater deviations from the beam's axis. (C) The target itself (a moth) reflects echoes that have relatively flat spectra except for the effects of overlap and interference between reflections from individual points (glints). The shaded plot shows 126 separate spectral slices in a waterfall display, made from successive incident sounds while the moth flutters it swings. Wingbeat periodicity is manifested by the horizontal bands in the plot. (D) The mild overall lowpass effect for echoes from the insect is due to greater incidence of destructive interference because wavelengths are shorter at higher frequencies. Echoes from realistic targets located at short ranges and on the broadcast beam readily can be distinguished from echoes returned by clutter located farther away or off-axis due to the pronounced lowpass filtering—specifically to the absence of strong FM2 and FM3 in clutter echoes.

From behavioral experiments, the bat's perception of shape is spatial in nature and depends on extracting the range separations of the target's principal glints from the interference spectrum of echoes. Crucially, the overall delay of an echo from a 1-glint target is registered on the same perceptual scale as the delay of either of the reflections from a 2-glint target. That is, the dual time- and frequency-domain sensitivity of cortical neurons creates a synthetic dimension of delay that the bat actually perceives. Targets such as insects being tracked by the bat will return several reflections that dominate the structure of echoes so that there are a small number of well-defined interference notches from which to derive estimates of the spacing of the glints. Objects located too far away or off-axis instead return echoes that have been altered more globally by lowpass filtering. The bat exploits these two extremes—the presence of a few well-defined glints with corresponding spectral notches for a target of interest *versus* a broad loss of frequencies that implies the presence of many poorly-defined glints for objects in the background—to suppress interference. There are clues that point to the use of small differences between the spectrograms of desired echoes and spectrograms of unwanted interference in how bats change their broadcast sounds to adapt to the presence of interfering sounds.

3 ADAPTIVE CHANGES IN BROADCASTS

During the course of interception maneuvers, bats change the duration, pulse interval, and frequency content of their sonar broadcasts. In the sequence shown in Figure 1, there is a gradual progression from longer (~3.5 ms) to shorter (0.5 ms), and for a progressive reduction in their 1st-harmonic frequencies from a tail-end sweep frequency of about 25 kHz at the beginning to about 16 kHz at the end, with commensurate sliding of all harmonics downward on the frequency scale. Sonar sounds in the first row of spectrograms contain 3 prominent harmonics, but the gradual reduction in FM1 frequency brings higher harmonics down into the overall frequency band, which extends up to 100-110 kHz for all of these sounds. Big brown bats capture flying prey by aerial maneuvers in open spaces as well as from substrates such as on or near vegetation, where echoes from clutter (unwanted objects surrounding the target) can create interference that might affect the bat's use of echolocation to find prey. Because these bats are so versatile in their foraging behavior, their echolocation has to adapt to different acoustic conditions—in open spaces, so-called "edge spaces," and also in locations even more closely surrounded by vegetation. The duration, interpulse interval, and frequency structure of broadcasts are changed from pulse to pulse according to the distance to the target being tracked or the distance to objects in the background, such as vegetation. In behavioral tests of echo-delay perception, as well as in aerial interceptions or obstacle-avoidance tests in the confined space of a flight room, broadcast durations typically are 2-5 ms. The presence of 3-6 harmonics in the sonar sounds of big brown bats is somewhat unusual for an aerial-feeding FM bat; many species that hunt for flying insects in open spaces typically cover the same wide frequency band in only 1 or 2 FM harmonics. The majority of FM bats that hunt in vegetation for prey, or for fruit or flowers to feed on pollen and nectar, emit sonar sounds with 3-5 harmonics. The presence of multiple harmonics in big brown bats sounds is an indication, confirmed by video observations, that these bats indeed sometimes do forage in vegetation, not just in open spaces or edge spaces.

Bats also change their transmitted sounds in response to the presence of other sounds in the environment, including the sounds emitted by other bats in the vicinity. Figure 3 shows spectrograms for echolocation sounds emitted by three big brown bats flying close to each other over as pond hunting for flying insects. The frequencies in these sounds—most conspicuously the tail-end (low) frequencies of the FM1 sweeps—differ from bat to bat by up to several kilohertz. It is widely assumed that bats of the same species shift the frequencies in their sounds away from each other to avoid interference from the sounds of other bats, although individual differences are present in big brown bats, too. The question of jamming avoidance is difficult to study in field recordings because the sounds of different bats moving in different directions would be Doppler-shifted by different amounts, and, furthermore, there is no way to know what each bat intends to do with the echoes it receives because their behavior is not controlled. To obtain quantitative evidence that shifts in broadcast frequencies indeed are made to avoid interference from other sounds, three big brown bats were trained in a two-choice target-detection task, and their signal frequencies measured while continuous pure-tone constant-frequency sounds were delivered as interference. The bats exhibited the same detection performance in the presence or absence of the interference. Each bat emitted sounds with slightly different tail-end "baseline" frequencies in the absence of interference (23.5, 25.7, 29.1 kHz; see Figure 4A), but they changed the frequencies of their broadcasts when the constant-frequency interfering sound encroached on the frequency region of their individual baseline frequencies. When the frequency of the interfering sound was below the bat's own preferred frequency by less than 1-4 kHz, the bat increased the frequencies of its broadcasts to raise the tail-end frequency away from the jamming sound. When the frequency of the interference was within 1-4 kHz above the bat's preferred frequency, the bat decreased its emitted frequencies to lower the tail-end frequency away from the jamming sound. The curves in Figure 4A show the abrupt reversal in the direction of the change in emitted frequency as the interfering tone came to coincide with the bat's baseline frequency. The signals that the bats emitted did not change in duration when the jamming sound changed in frequency (Figure 4B), so the frequency changes appear to be a jamming avoidance response applied to emitted frequency itself, not just a lowering or raising of frequency that occurs incidental to changes in duration.

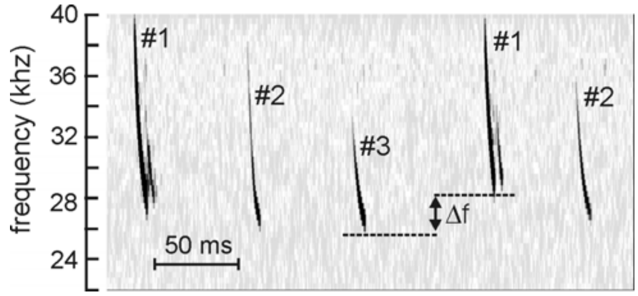

Figure 3: Interaction among bats flying near each other

These spectrograms show the low-frequency end of FM1 sweeps in sonar sounds emitted by three different big brown bats flying and searching for insects over a pond. These sounds differ in frequency at the tail-end of the FM1 sweep. Frequency differences (Δf) between bats flying together usually are up to 2-5 kHz.

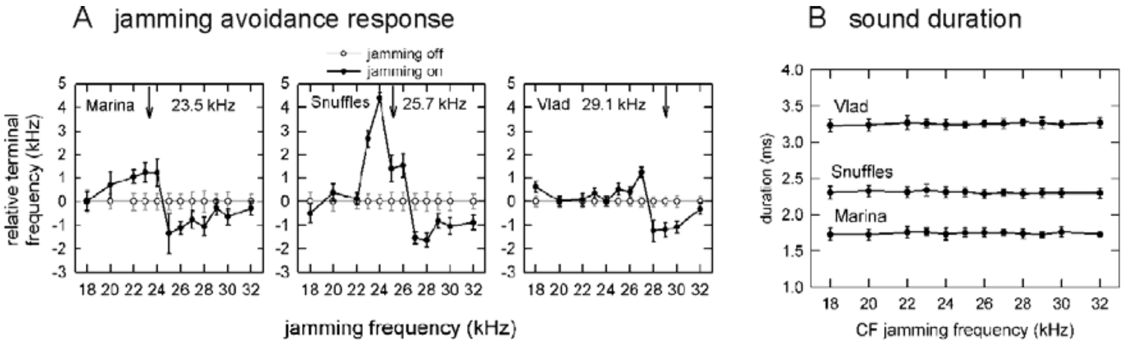

Figure 4: Jamming avoidance response of three bats

(A) Relative frequency shifts (mean ± 99% confidence intervals) of terminal FM1 frequencies emitted by three big brown bats (vertical axis) during target detection task in presence or absence of jamming sounds at different constant frequencies (horizontal axis). Mean absolute tail-end "baseline" frequency without jamming is given on the plot for each bat. (B) Durations of emitted sounds (mean ± 99% confidence intervals) for same sounds as shown in A. The bats shifted the frequency at the tail-end of the FM1 sweep according to the frequency of the jamming sounds without changing the duration of the sounds.

Echoes returned from background objects—clutter—are perhaps the most prevalent and potentially the most serious source of FM sounds that could cause interference for the bat's soanr. Echoes reflected by vegetation or the ground necessarily resemble echoes from targets of interest because they are all reflections of the same broadcasts. Moreover, in many instances parts of the clutter itself constitute the "target of interest," such as when a bat flies through vegetation and has to repeatedly avoid colliding with the branches immediately to its front. Judging from video recordings of bats flying through vegetation in the dark, bats must be very good at locating and dodging around obstacles while also being sufficiently aware of objects off to the side or farther away to plan their flight paths accordingly.

To investigate the behavior of bats flying in dense, range-extended clutter, a new procedure was devised that challenges the bat with a complex obstacle field while still allowing for quantitative assessment of the bat's reactions. Figure 5 illustrates the experimental design. The obstacles themselves are the first innovation: To replicate the density and depth of obstacles encountered by bats flying through vegetation, a flight room was set up with an array of chains hanging vertically from ceiling to floor, with successive rows of chains spaced at intervals of 40 cm over a total extent of 4 m to almost 6 m. To provide a path through the chain array, individual chains were removed from the array to give the bat an alley 75-100 cm wide for the bat to fly along. The critical design feature is that, when the bat enters the array at the start of the alley, it receives large numbers of echoes from chains in the array, at distances from as close as the nearest chains to the left and right, and as distant as the last row of chains. Consequently, each time the bat emits a sonar sound, clusters of echoes return at intervals of about 2.3-2.5 ms from row after row of chains for 30 ms or more (Figure 5A). The chains themselves are very strong targets because each link is large, and attenuation of echoes from an individual chain amounts to only about 11-18 dB as distance increases from 30 cm to 200 cm, partly because, as the distance increases, progressively more of the links in the chain are recruited into the beam of the incident sound. Figure 6B,D shows a practical acoustic demonstration of the effectiveness of the chains as obstacles through recordings of the chain echoes from the bat's own sounds. During each trial, the bat's flight-path is traced for 3D reconstruction by making stereo-frame video recordings with two thermal-imaging video cameras (to allow flights in total darkness), digitizing the location of the bat in each frame, and using video motion-analysis software to build 3D data sets containing the bat's flight path and the locations of all the chains (Figure 5B). The position of the bat each time it emits a sonar sound is part of the data-set, too. The most crucial methodological innovation in these chain-flight experiments is use of a radio microphone (the "Telemike") to provide "on-board" recordings of the bat's sonar sounds while the bat is flying (see Figure 5A). Recordings of the broadcasts and echoes from a bat flying in the chain array while using the Telemike are illustrated in Figure 6. The Telemike's advantage is that it gives an unbiased recording of the bat's sounds from the on-board microphone located between the bat's ears, without the Doppler shifts that occur in recordings made with a microphone remote to the bat. A second ultrasonic microphone, more sensitive than the Telemike unit, is placed behind the bat and aimed into the chain array to pick up echoes from the chains while the bat is in flight. The chain array was intended to be deep enough to yield echoes for at least 20-30 ms following each broadcast, and yet to be sufficiently complex and dense as to require the bat to emit sounds so rapidly (to negotiate the array) that echoes from one broadcast would still be arriving when the bat emitted its next sound. This intention was borne out in the bat's behavior.

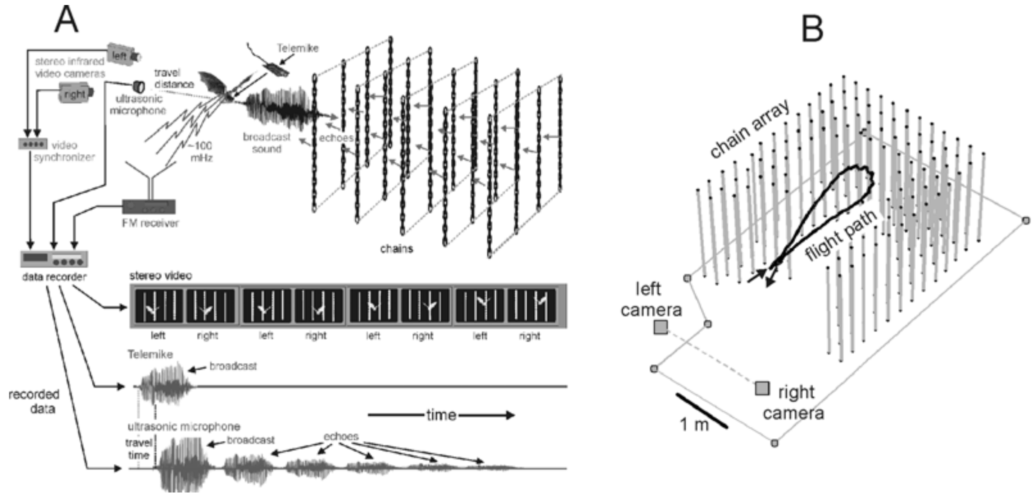

Figure 5: Diagram of obstacle-avoidance tests with the chain array and Telemike

(A) Big brown bats are trained to fly along a flight room and land on the far wall after negotiating an obstacle array of vertically hanging chains [30]. The bat's ultrasonic sounds are picked up by an "on-board" radiomicrophone (Telemike) that transmits the bat's sounds on a radio carrier (~100 MHz) to an FM radio receiver, which demodulates the sounds from the carrier and records them on a digital data recorder. Two synchronized thermal-imaging infrared cameras provide stereo video images of the bat's flight at 30 frames/s that are merged into a single video stream and recorded on the same Sony. Echoes reflected back to the bat by the chains are recorded with a sensitive ultrasonic microphone placed behind the bat as it flies towards the chains and pointing along its flight path. (B) 3D reconstruction showing one flight path (chains are vertical gray lines), with the bat approaching the chains by flying into a narrow open area within the array and circling to fly back out. The bat's acoustic behavior when it flies into this open space while receiving echoes from all of the rows of chains is shown in Figure 5.

When flying through the chains, big brown bats emit their sonar sounds in groups spaced 20-25 ms apart, mostly pairs, called "strobe groups," separated by longer intervals of about 50 ms (Figure 6A). The strobe groups themselves are far enough apart for echoes from the chains to have completed their arrivals (*i.e.*, the echo-receiving epoch is finished) before the next sound is emitted. However, for sounds *within* the strobe group, echoes from the first sound still are being received (the echo epoch is still in progress) when the second sound is emitted (Figure 6B). Because echoes from different broadcasts are mixed together in the interval following the second sound in the strobe group, ambiguity exists about how to interpret the arrival-times of echoes for estimating target range. The flying bat reacts to this ambiguity by raising the tail-end frequency of the first sound in the strobe group and lowering the tail-end frequency of the second sound, creating a frequency difference (Δf in Figure 5C) that "tags" the echoes of the first sound as being different from the echoes of the second sound by this frequency difference (Figure 6D). The size of the frequency shift (Δf) is up to about 5-8 kHz. Figure 6D illustrates how the spectrograms of the echoes of the two sounds in the strobe group (*e.g.*, #3 and #4 in Figure 6D) become readily separable into two different "auditory streams" as a consequence of the frequency shift, which the Telemike reveals to be made by the bat, not caused by different amounts of Doppler shift. That this is a deliberate reaction to resist interference from clutter is demonstrated by what happens when the bat has flown far enough into the chain array that echoes from successive chains no longer continue to arrive over an epoch of 20-30 ms. If the echo epoch is short enough that echoes of the first sound in the strobe group are over before the second sound is emitted, then the bat stops making the frequency difference (Δf), so that both sounds now contain the same tail-end frequencies (*e.g.*, #14 and #15 in Figure 6D).

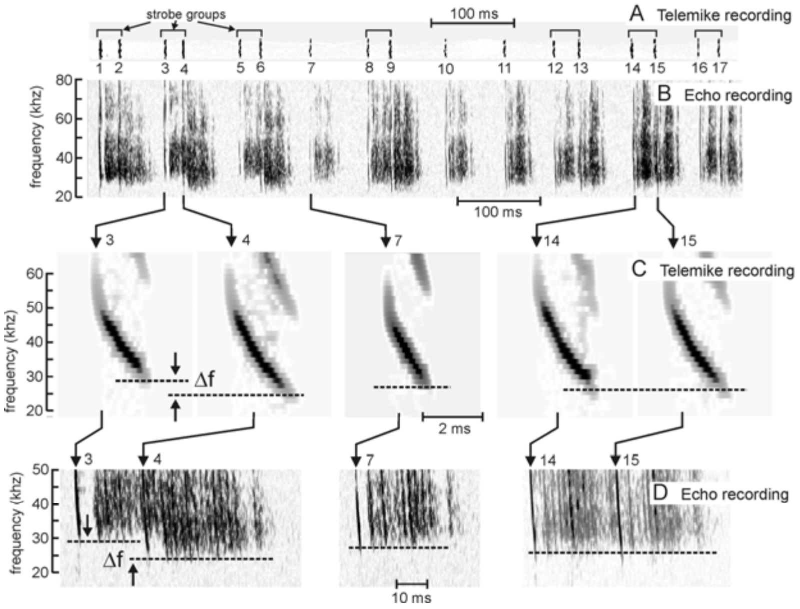

Figure 6: The big brown bat's adaptive response to dense clutter

(A) Spectrograms for a sequence of 17 sonar sounds emitted during the flight shown in Figure 4B, to indicate the timing of the broadcasts and the grouping of sounds into "strobe groups" [4,23]. These signals were recorded with the Telemike. (B) Spectrograms for the same sequence of broadcasts picked up instead with the ultrasonic microphone used to record echoes from the chains. The dark smear following each broadcast is the epoch of time during which echoes from the chains are returned to the microphone (and to the bat). (C) Expanded spectrograms for five numbered broadcasts recorded by the Telemike. The on-board microphone acquired the sounds without being influenced by Doppler shifts, so the frequency difference (Δf = 7 kHz) between sounds #3 and 4 in the same strobe group is generated by the bat itself. (D) Expanded spectrograms for the ultrasonic echo recordings of the same five numbered broadcasts to show the projection of the frequency difference (Δf) onto the echoes of the corresponding broadcasts. When the echo epochs overlap (#3 and #4), the bat separates the broadcasts and echoes by spreading their tail-end FM1 frequencies apart. When the echo epochs do not overlap, the frequency spreading is absent (#7 is a single sound, while for sounds #14 and #15 the bat has flown far enough into the chain array that the echo epoch is shorter).

4 DELAY PERCEPTION FOR HARMONICALLY ALTERED SPECTROGRAMS

The bat's strategy for streaming the sounds and their corresponding echoes in the strobe groups while flying through the chains (Figure 6) looks similar to its behavior when flying near other bats in the field (Figure 3) or when avoiding jamming from pure-tone interfering sounds (Figure 4). Changes in tail-end frequencies of only a few kiloHertz evidently are sufficient to render the echoes of these sounds distinct from echoes of other sounds, and, by extension, to render the sounds of other bats flying nearby as different enough that they are not confused with actual echoes. Added to these marginal (mostly <3-8 kHz out of ~80-kHz bandwidths) changes in broadcast frequencies is the sound-to-sound variability in spectrograms of broadcasts emitted during interceptions (Figure 1). These frequency changes actually have an impact on the entire spectrogram through the changes they cause in the slope, the sweep shape, frequency range, and the harmonic structure of FM sweeps in the bat's sounds. To understand what bats perceive when they receive FM sounds that contain slight differences in their frequencies or differences in their FM sweeps and harmonic content requires new psychophysical experiments that differ in several respects from the types of target and echo discrimination experiments used in the past. For present purposes we will concentrate on the role of the harmonics (FM1, FM2) and the effects of the lowpass filtering imposed on echoes by a target's off-axis location or excessive range (Figure 2).

Figure 7 shows a diagram of a psychophysical experiment designed to investigate how changing the spectrogram of echoes affects perception of echo delay. The psychophysical method is two-alternative forced-choice (2AFC), or 2-choice simultaneous discrimination of rewarded echoes (S+) at a delay of 3160 μs and unrewarded echoes (S-) at a delay of 3960 μs. The bat's response is to move forward towards the loudspeaker that produces the shorter delay (for S+) to receive its food reward (piece of mealworm). The basic method is to train each bat to sit on the elevated Y-shaped platform and broadcast its sonar sounds into microphones (m) located on the left and right. Sounds picked up by the microphones are delayed electronically, and then filtered to deliver either FM1+FM2 normally, or to remove either FM1 or FM2 so that echoes would contain only one harmonic, and returned to the bat as echoes from loudspeakers (s). The rewarded (S+) and unrewarded (S-) stimuli were presented on the bat's left or right in randomized sequences over multiple trials. Figure 8 shows spectrograms determined from electronic filter settings for the three different stimulus conditions (FM1+FM2, FM1 alone, or FM2 alone). Once the bat is trained and can perform the task for the normal echoes (FM1+FM2 in Figure 8) at better than 90% correct responses, sets of 50 experimental trials were conducted each day for three days before changing to either of the filtered echo stimuli (FM1 alone or FM2 alone in Figure 8). For each stimulus condition, 150 trials were carried out on each bat, for a total of 600 trials on each of the stimulus conditions (FM1+FM2, FM1 alone, or FM2 alone).

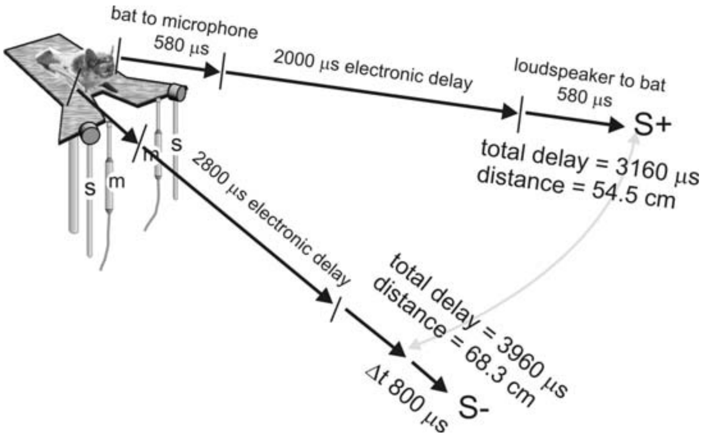

Figure 7: Experimental method for measuring perception of echo delay

Diagram of the electronic target simulator that presents bats with echoes arriving at delays of 3160 μs (rewarded stimulus, or S+) or 3960 μs (unrewarded stimulus, or S-). Difference in delay is 800 μs. S+ and S- were randomly alternated from left to right on successive trials. The experiment measures the performance of bats (percent errors) on this 800-μs task when echoes contain both FM1 and FM2, or when either FM1 or FM2 is removed from echoes (see Figure 8).

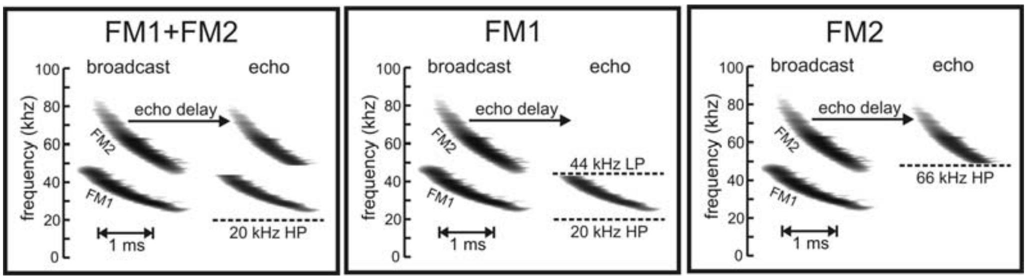

Figure 8: Spectrograms of echoes in delay discrimination tests

Spectrograms for echoes delivered to the bat as S+ and S- in the experiment shown in Figure 7. The harmonic content of echoes is determined by highpass and lowpass filters inserted into the electronic simulator.

The results of the experiment are illustrated in Figure 9. All four bats behaved similarly in the experiments, with performance being very high (>90% correct) for the normal echoes containing FM1+FM2, and only slightly lower (~87% correct) for echoes containing only FM1. However, for echoes containing only FM2, performance was very close to chance (47-57% correct). The bats clearly treated echoes containing FM1 and FM2 about the same as echoes containing only FM1. In both cases they were able to assess the delay of echoes effectively. In contrast, they were not able to assess the delay of echoes containing only FM2. Recall that, from Figure 2, echoes from targets located on-axis and at close range will contain both FM1 and FM2 naturally, while echoes from targets in the background will contain only FM1 at substantial strength. In normal acoustic conditions, echoes that contain only FM2 are not ever likely to occur. The asymmetry illustrated in Figure 2 for the presence of FM1 and FM2 in naturally-occurring echoes from targets located at different positions and different distances around the bat is reproduced in the bat’s perception of delays for echoes containing different combinations of FM1 and FM2.

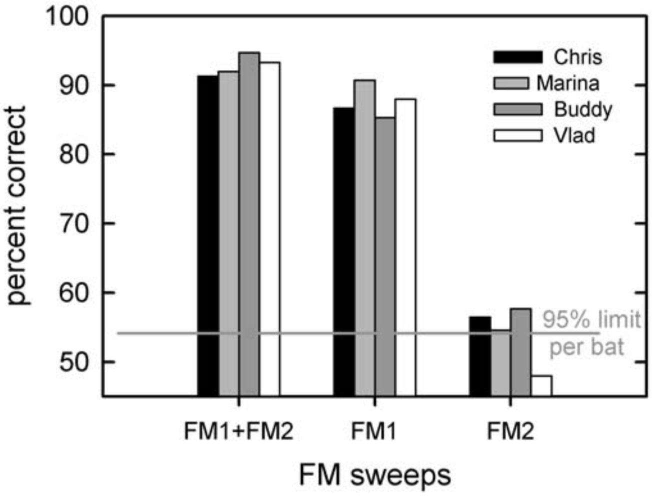

Figure 9: Percent correct responses

Performance of four bats on 150 trials each in the three stimulus conditions shown in Figure 8, for a total of 600 trials per condition. Perfect performance is 100% correct, and chance performance is

50% correct. The horizontal gray line shows the level of performance above chance that corresponds to the 95% confidence limit for a single bat in each condition of the experiment.

5 CONCLUSIONS

The big brown bat’s sonar sounds are beamed to the front with a directivity that depends on frequency. At frequencies of 20-30 kHz, which appear at the low end of FM1, the emitted beam is very broad, but at higher frequencies the beam becomes progressively narrower, so that frequencies of 55-90 kHz in FM2 are localized in a narrower zone in the center of the beam. Figure 2 illustrates how the frequency dependence of emission directionality causes different amounts of FM2 to be present at different locations in the beam, while the more broadly directed FM1 is present more nearly uniformly in these same locations. The signal that impinges on targets located directly on the axis of the broadcast beam thus contains both FM1 and FM2, so that echoes returning to the bat can convey information about the target’s shape across the entire bandwidth of the sounds. However, at locations off to the left or right of the axis, the strength of FM2 is reduced compared to the strength of FM1, so that, in effect, the sound impinging on targets contains primarily FM1 alone. A similar effect occurs for incident sounds projected outward to greater distances from the bat. Atmospheric attenuation is 0.4-0.7 dB/m at frequencies of 20-30 kHz in FM1, but attenuation becomes progressively greater at higher frequencies, rising to 2-3 dB/m at 50-90 kHz in FM2. Thus, the bat sound that arrives to impinge on targets at greater distances contains less FM2 compared to FM1. This selective removal of FM2 occurs regardless of the direction of the target, even for locations on the axis of the beam. Targets located both off the axis and farther away receive an incident sound that contains no FM2 at all. Except for interference caused by multiple overlapping reflections from different body-parts, insect-sized targets reflect echoes at frequencies of 20-100 kHz with about equal strength, so the effects of the target itself are faithfully returned to the bat from locations near the center of the broadcast beam and within distances of 1-2 m from the bat. For echolocation, this zone appears to function in a manner analogous to the fovea in vision; objects are perceived with high accuracy if they are located in this zone. Targets located in this zone are ensonified with both FM1 and FM2, and the bat uses auditory computations based on the harmonic structure of echoes to create the equivalent of a sonar “fovea” for perceiving the structure of objects. Big brown bats can accurately judge the delay of electronic echoes containing both FM1 and FM2, which would apply to echoes from real targets in the zone located directly to the front and within 1-2 m. For targets located outside of this zone—off the axis of the beam or farther away than 1-2 m, the low-pass filtering effects of broadcast directionality or atmospheric absorption cause echoes to contain mostly FM1. The bat’s delay acuity in the experiment for echoes containing just FM1 is only slightly reduced compared to echoes containing FM1 and FM2. The necessarily lowpass effects of the bat’s acoustic environment do not result in selective removal of FM1 compared to FM2, however, so that bats would never have to process naturally-occurring echoes containing only FM2. (No high-pass filtering occurs naturally in echolocation.) The discrimination performance of the bats is very poor for echoes containing only FM2. In fact, it is not clear whether the bats even treated stimuli containing only FM2 as echoes at all. Overall, the results suggest that the bat’s auditory system is organized to recognize only echoes whose spectrum reflects naturally-occurring acoustic low-pass effects of beaming and propagation combined with the effects of the target itself. Unnatural echoes containing only FM2 are either rejected or are processed to determine delay with very low accuracy.

ACKNOWLEDGEMENTS

These experiments were supported by ONR Grant # N00014-04-I-0415, by NIH Grant #R01-MH069633, by the Rhode Island Space Grant from NASA, and by JSPS Grant #S07704.