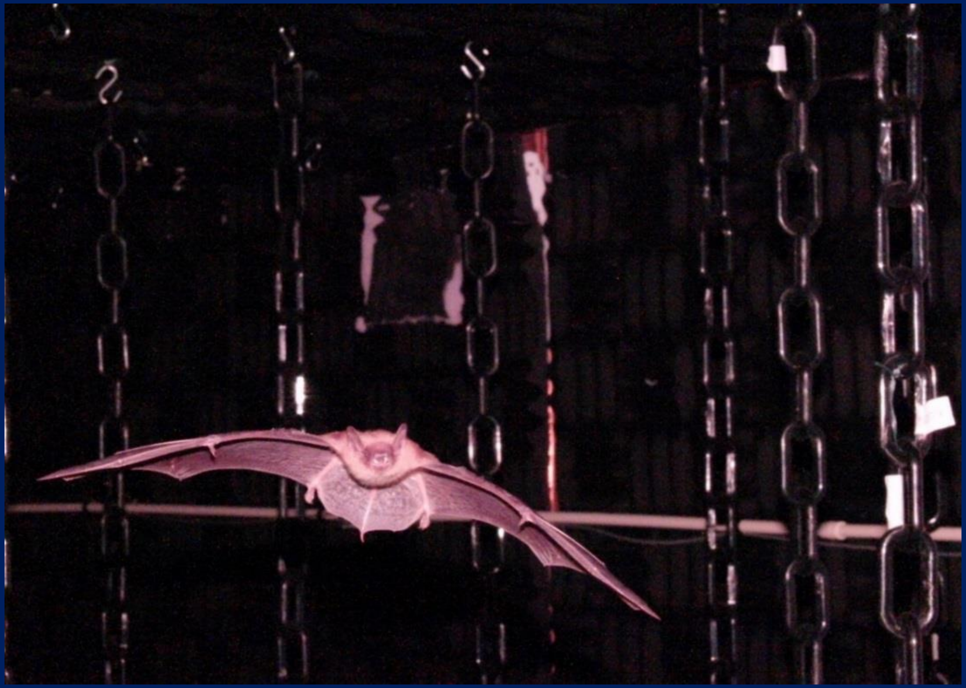
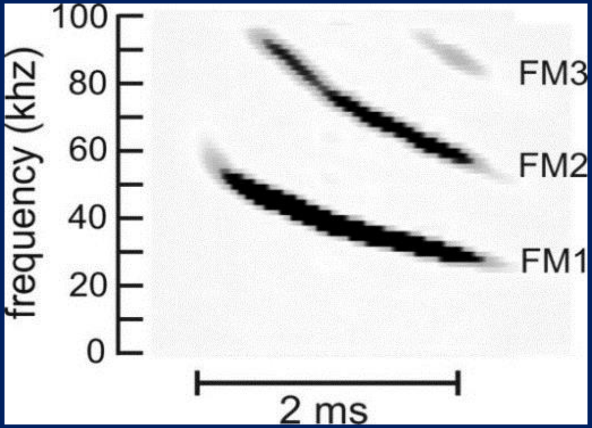
Time- and frequency- domain convergence onto a unified perceptual dimension in hearing

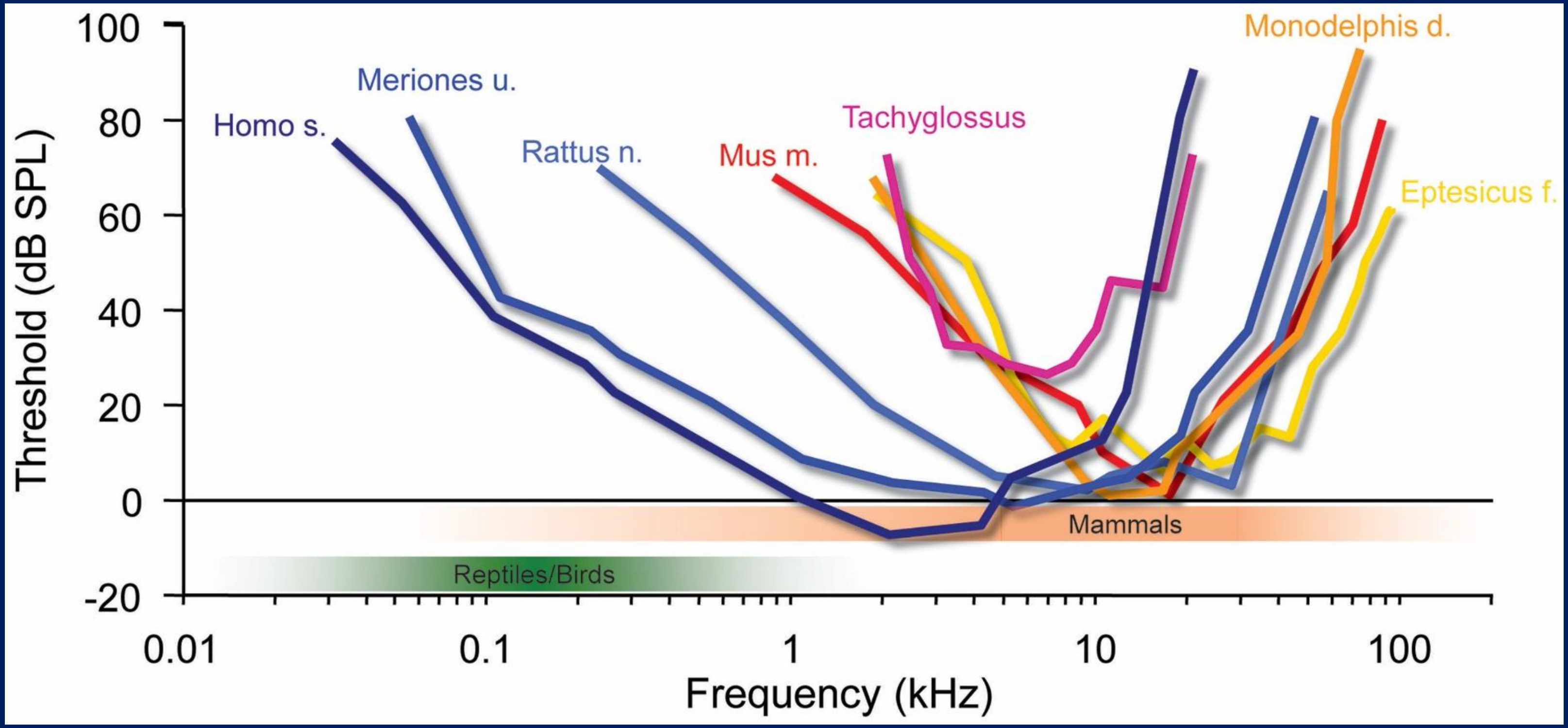
James A. Simmons
Department of Neuroscience
Brown University
Providence, RI 02912 USA
james_simmons@brown.edu

July 14, 2017
KITP Physics of Hearing

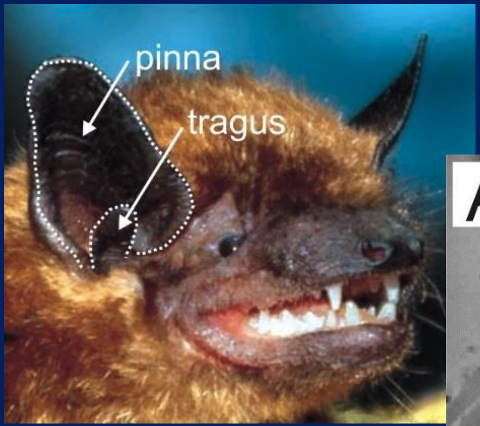


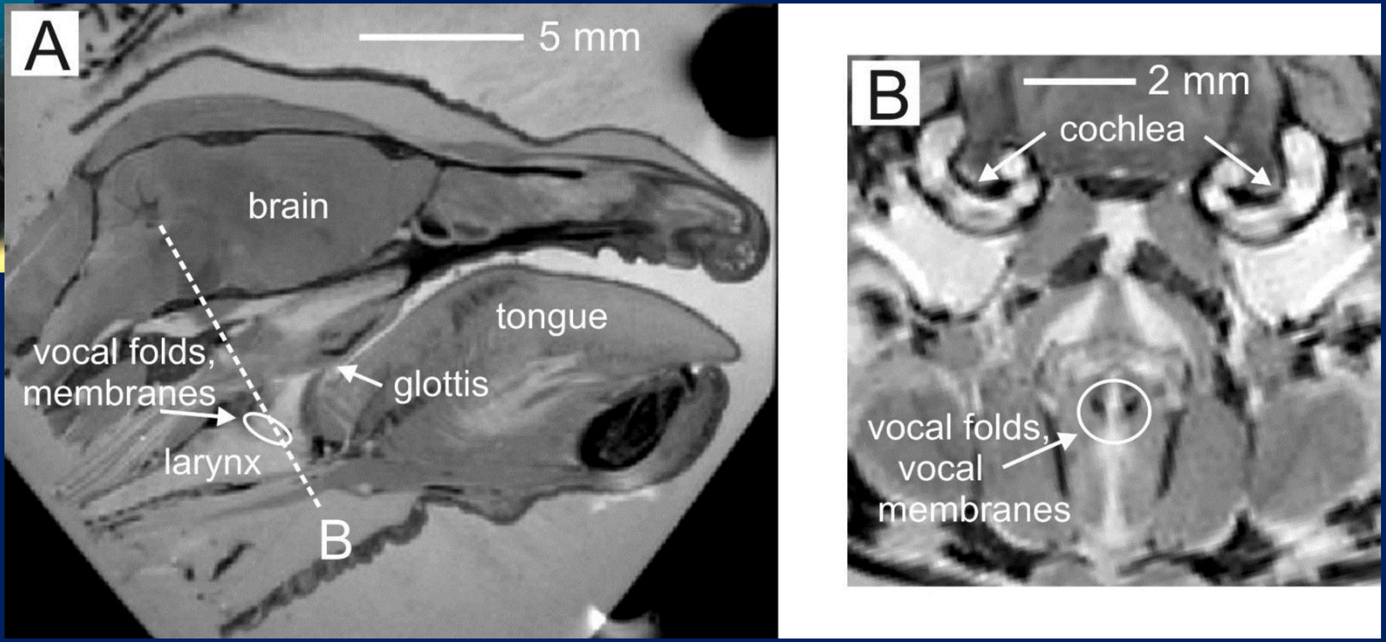


The big brown bat (*Eptesicus fuscus*) transmits FM biosonar sounds covering frequencies from 20 to 100 kHz in two or more down-sweeping harmonics (FM1, FM2). These signals are wideband, with a bandwidth of 75-80 kHz around a center frequency of 60 kHz.



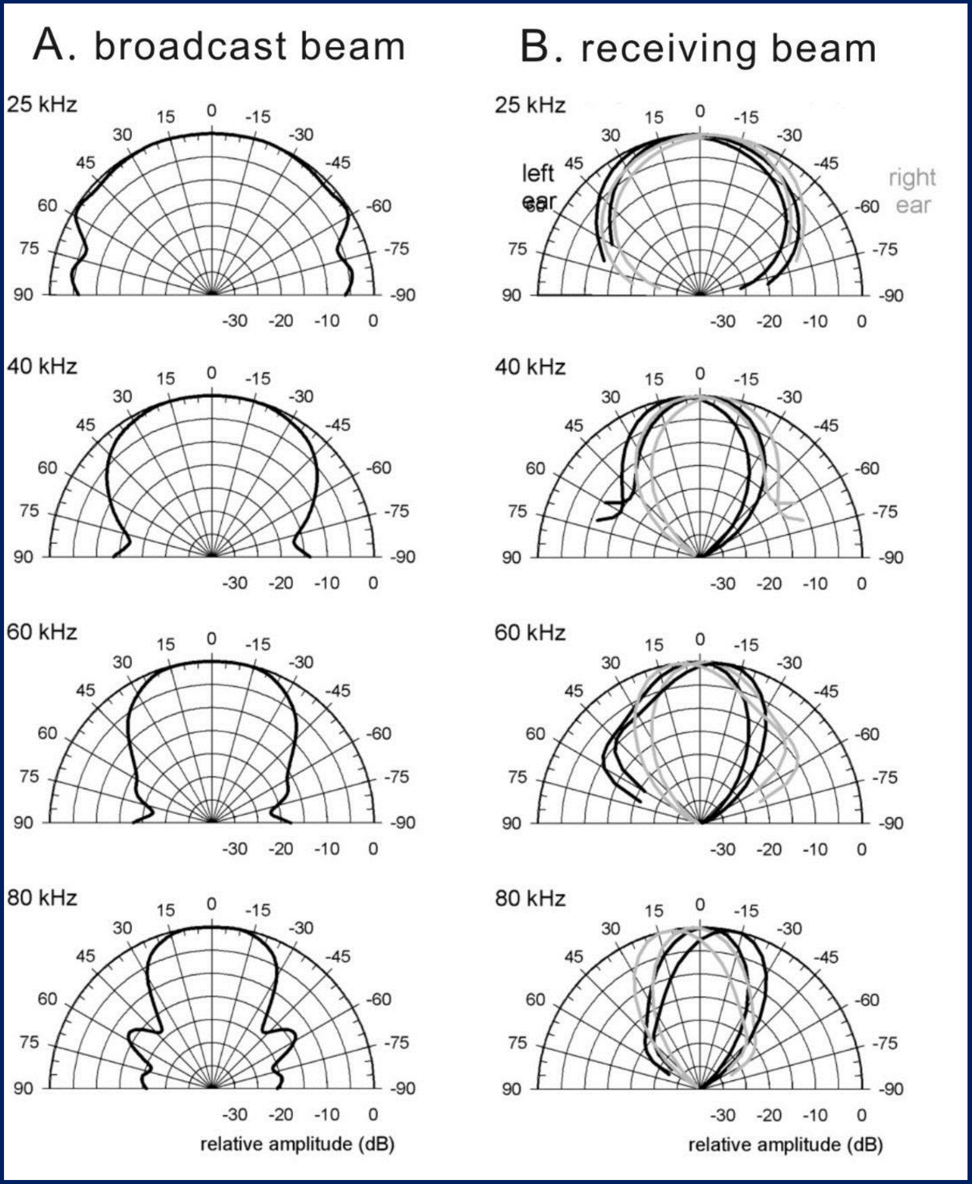
The big brown bat (*Eptesicus fuscus*) hears sound frequencies from about 5 kHz to 90-100 kHz and has outer-hair-cell active amplification throughout its hearing range.

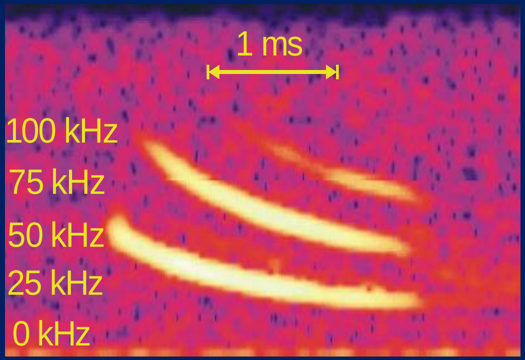


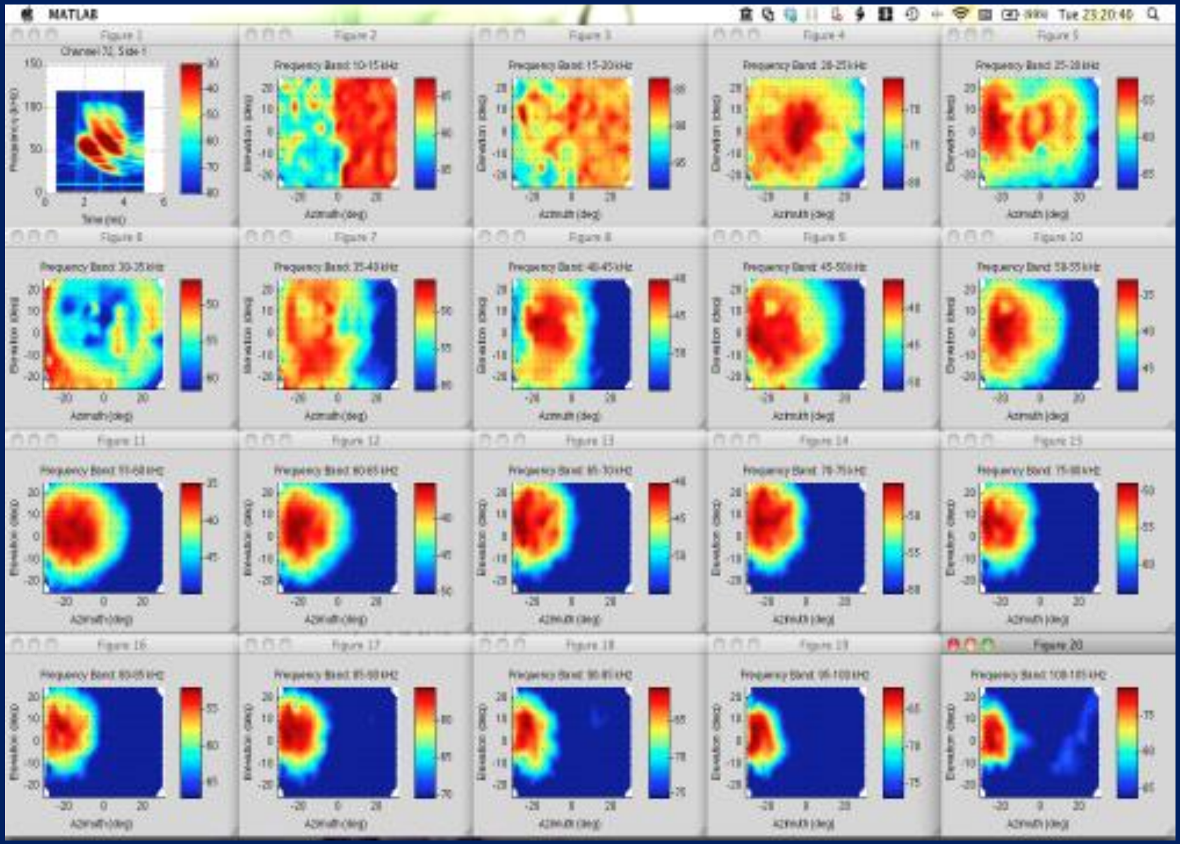


The bat's larynx and vocal tract act as the transmitter. Each sound is a rapidly downward-sweeping vowel. There are two receiving organs—the left and right inner ears—that form a two-element spatial array for acquiring directionally-dependent time differences.

Sounds are emitted through the open mouth. The bat's broadcasting and receiving beams are wide. Objects all around the bat are ensonified and their echoes are picked up by the ears.

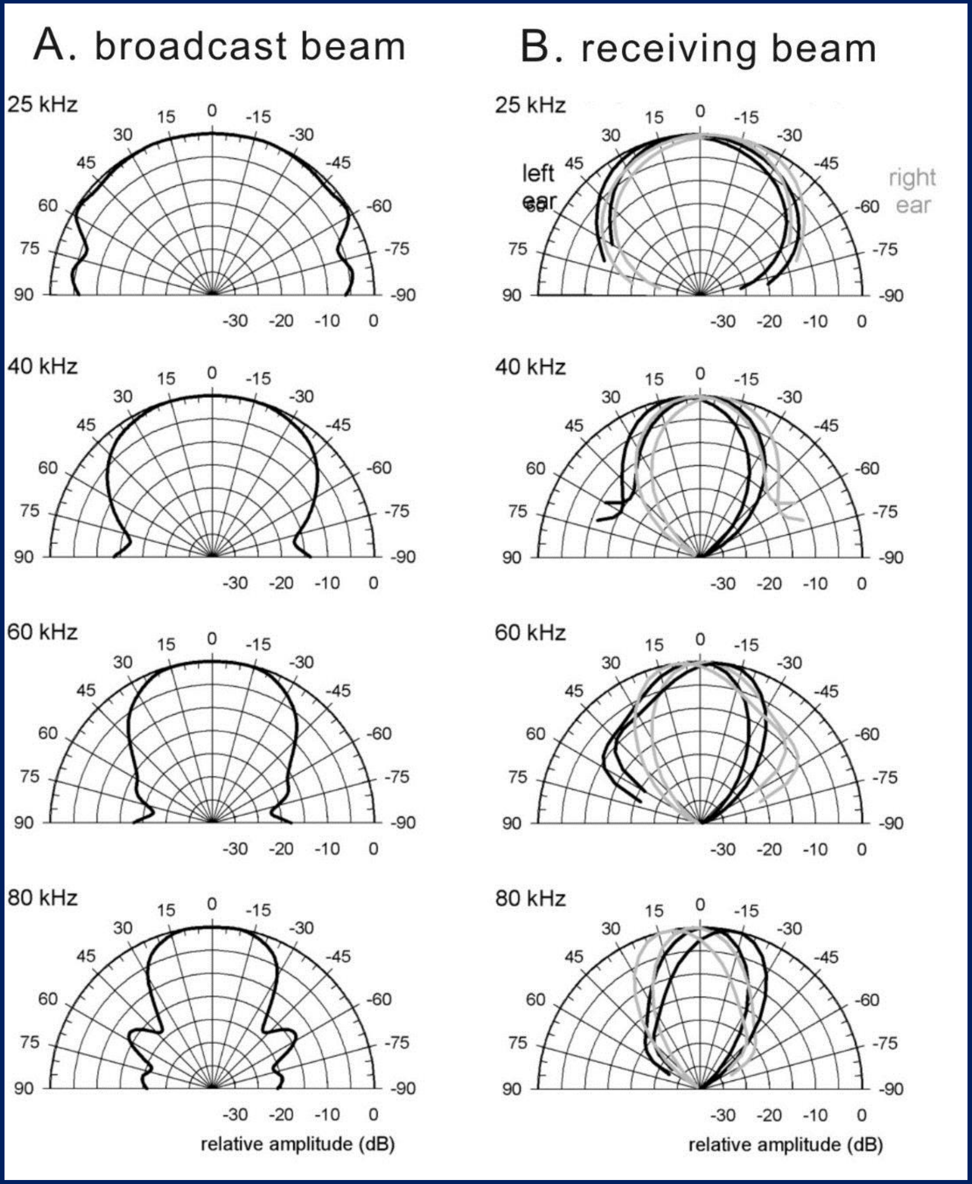




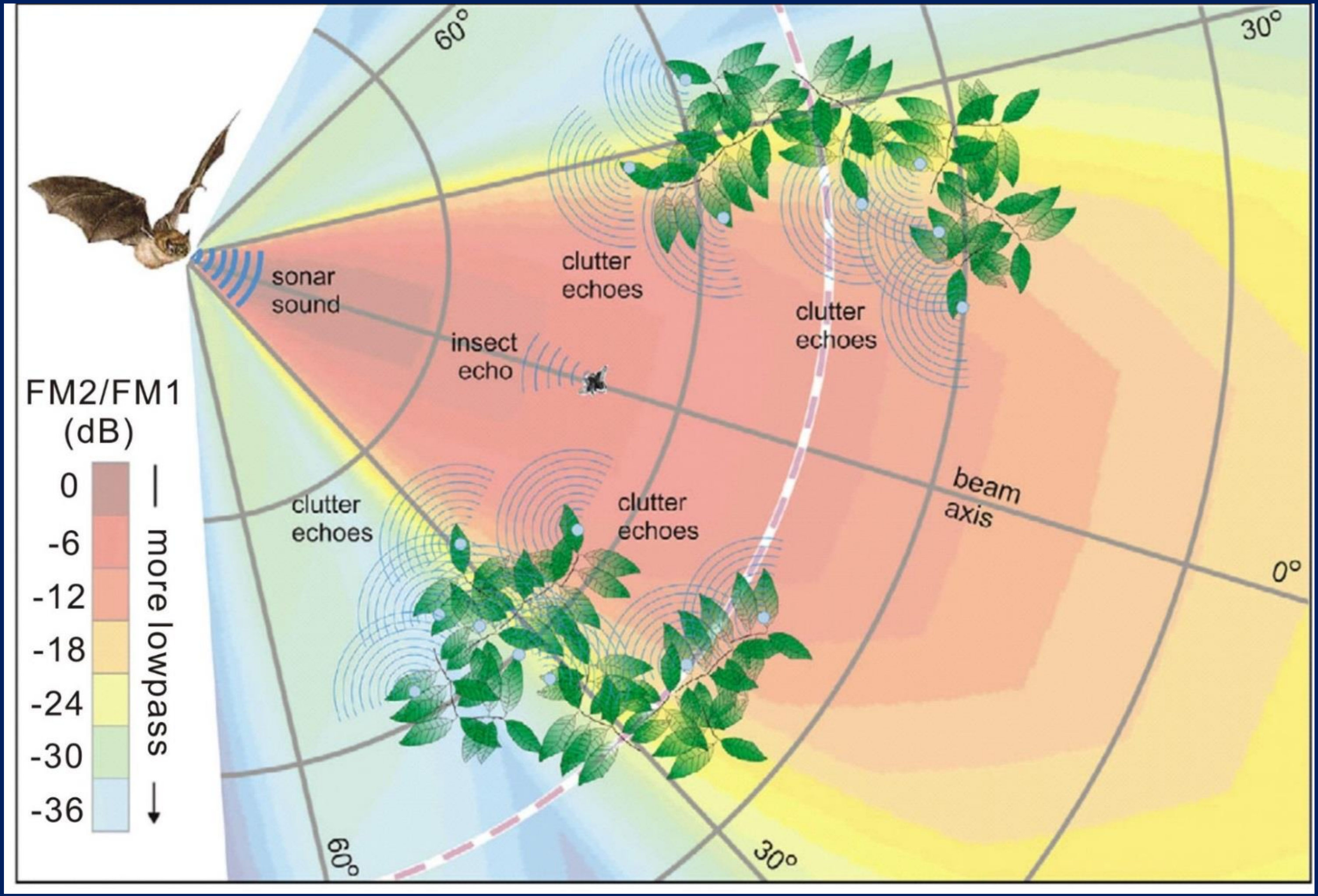


The beam is very broad at frequencies in FM1 and less broad at frequencies in FM2.

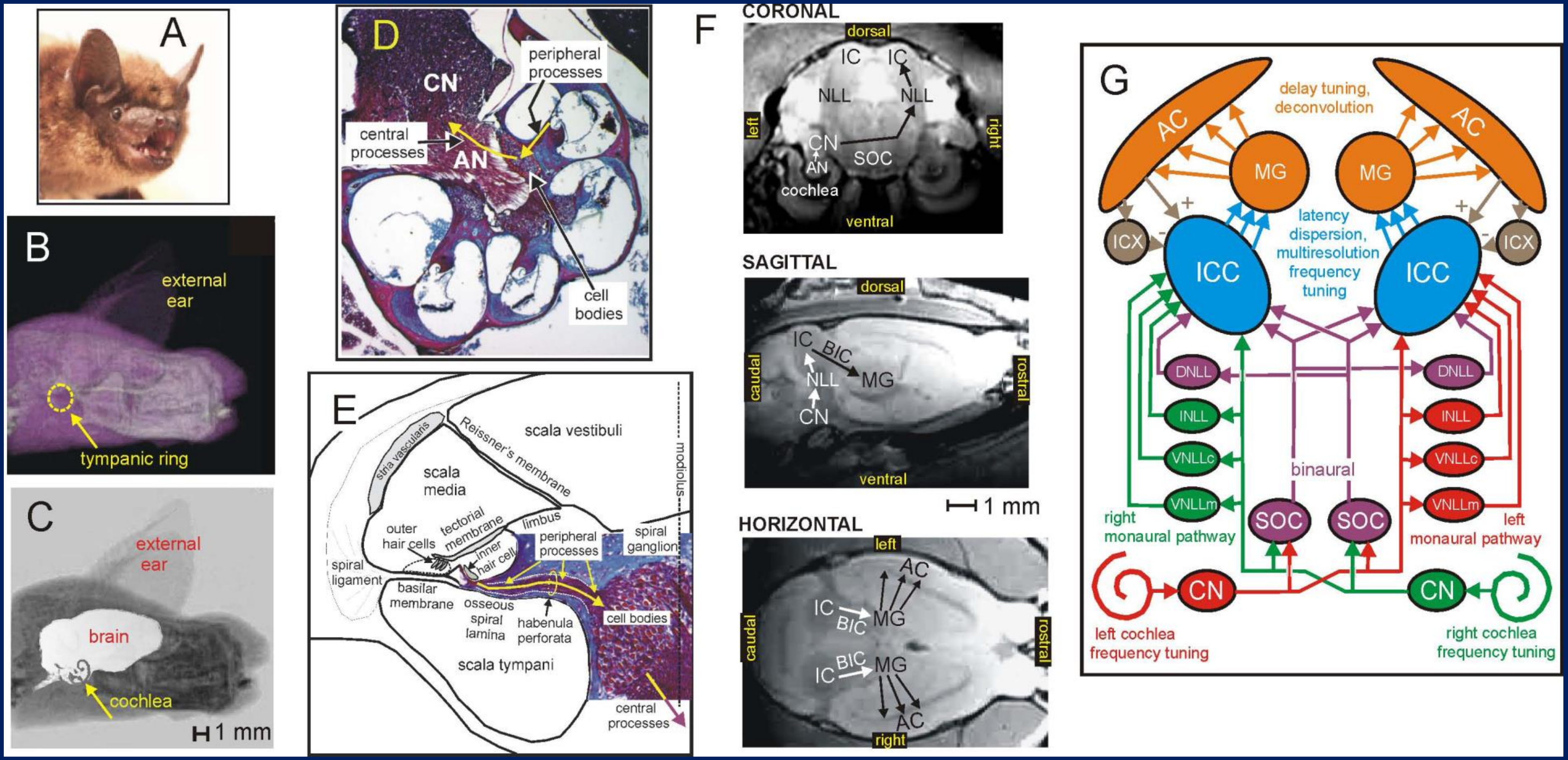
40 degrees



Crucially, only the full bandwidth of the sounds is emitted and received from straight ahead, on the acoustic axis of the system. Off-side or distant objects return lowpass echoes.

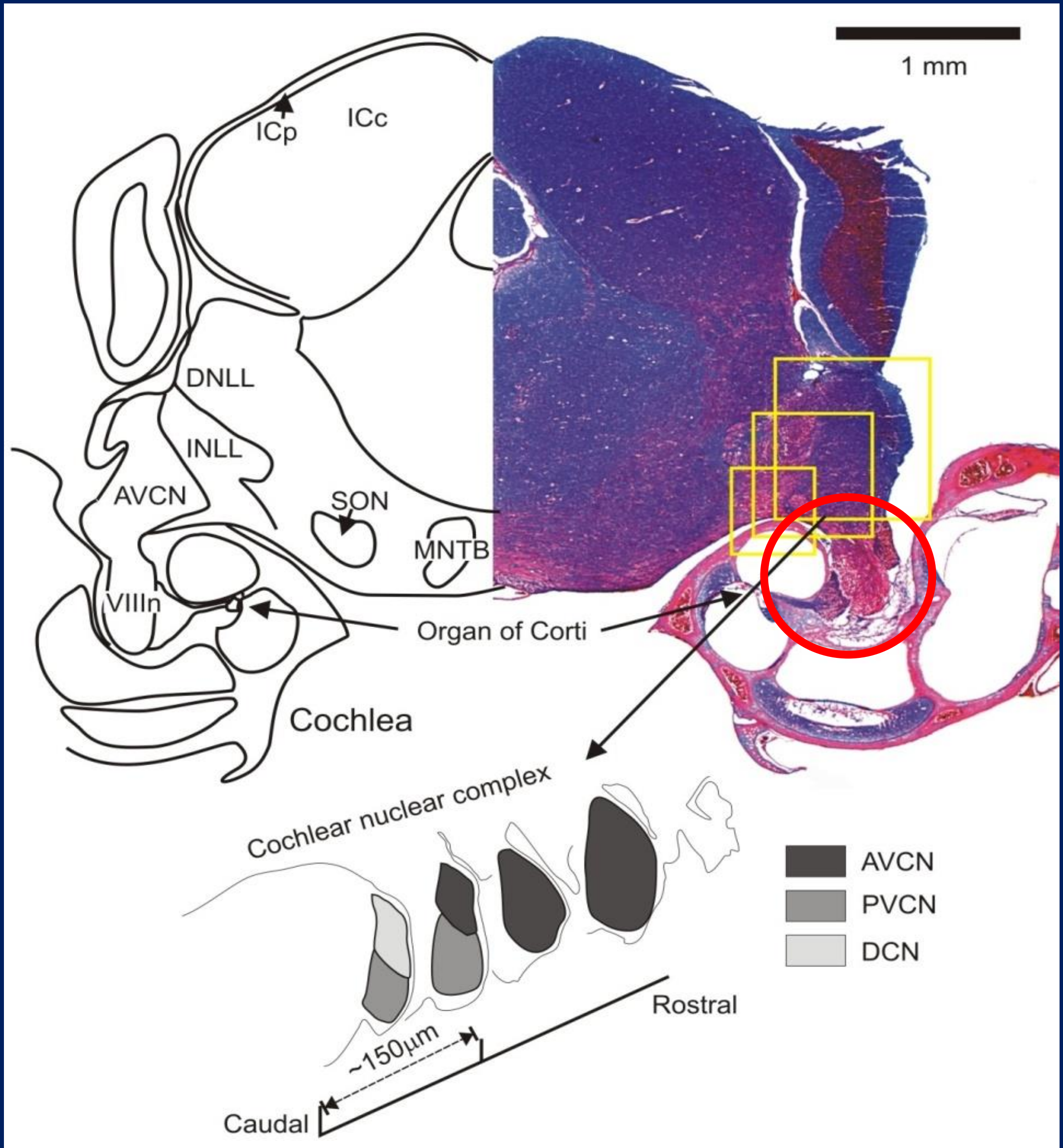


Acoustics of a biosonar scene.

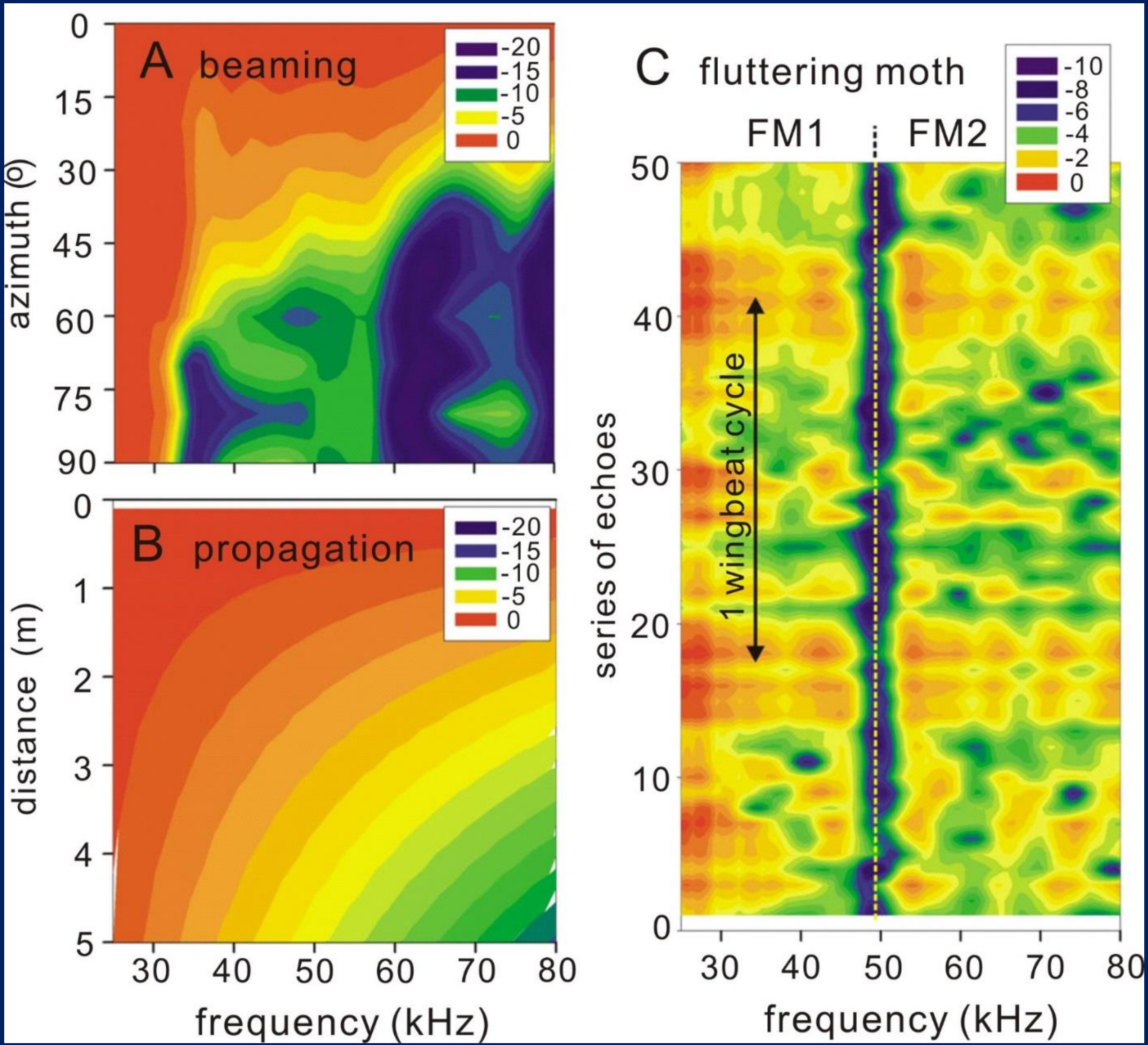


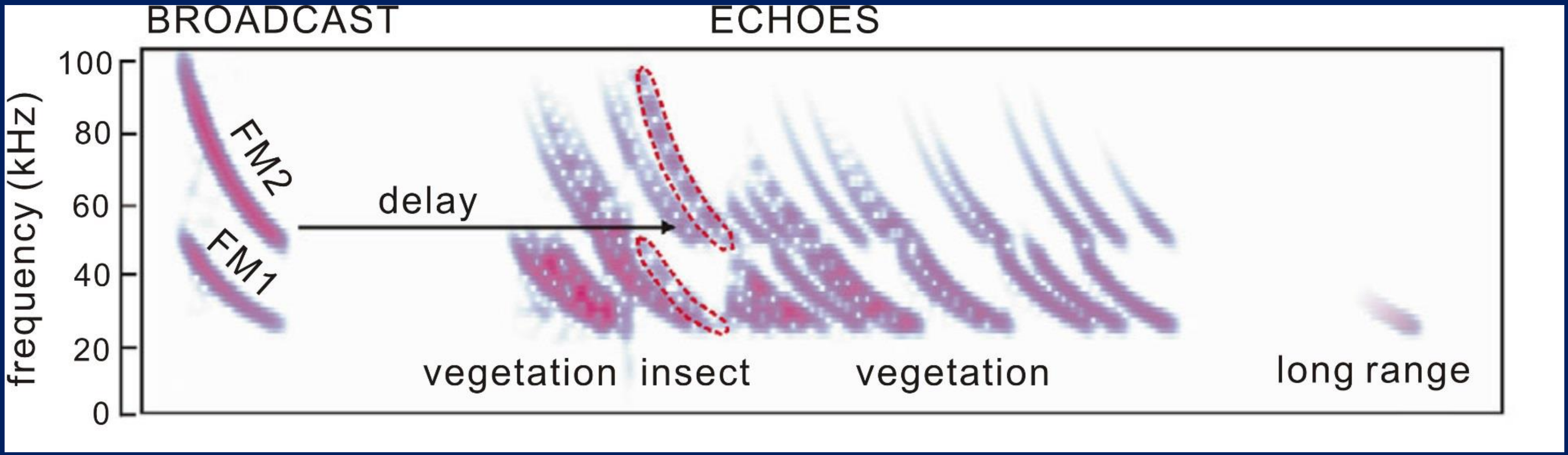
Both broadcasts and echoes are received by the inner ears, registered as a series of frequencies in the FM sweeps, and then marked by single spikes in neurons tuned to these frequencies all along the bat's auditory pathways. The representation is instantaneous frequency, with a twist related to phase.

The bat's cochlea is located very close to the brain, so the auditory nerve is short, which may minimize the internal noise of the transduction process. Surprisingly, big brown bats perceive the phase of FM echoes as a component of echo delay.



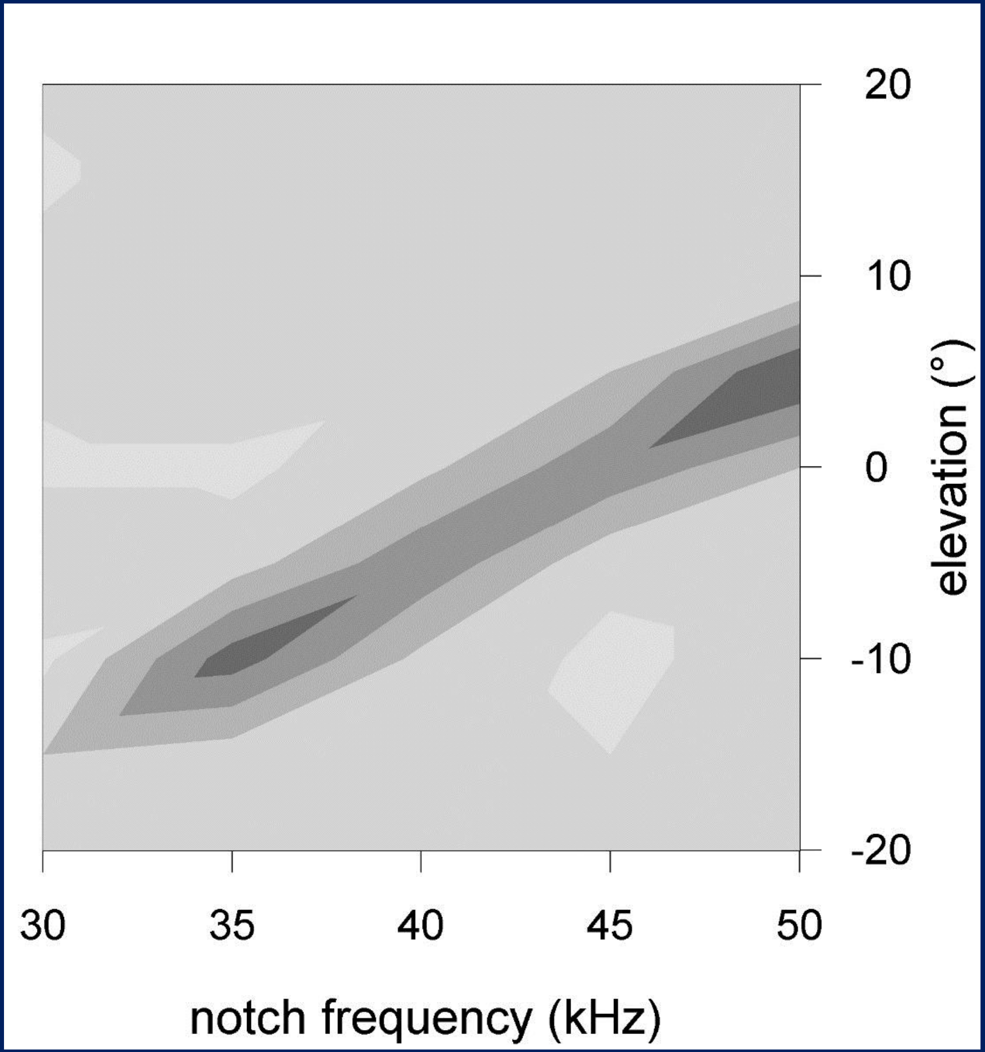
The acoustic effects of location in the scene (A, direction; B, distance) are lowpass in character. Acoustic features for objects in the axis of the sonar beam (C) are wideband with spectral ripple that shifts according to the target's aspect angle and fluttering motions.

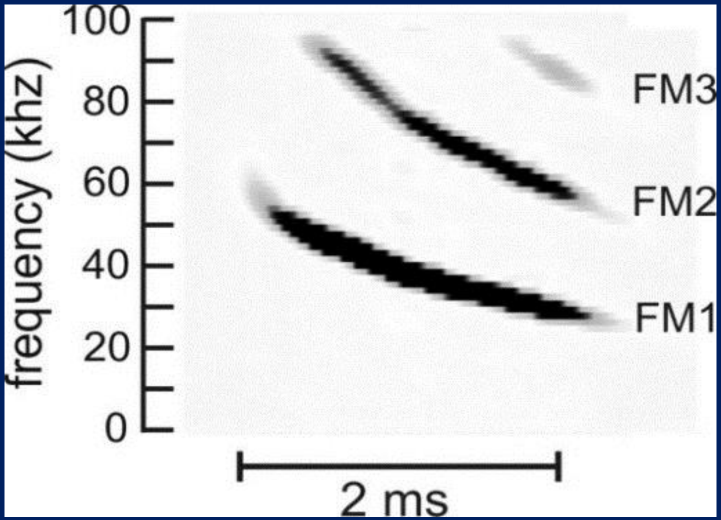


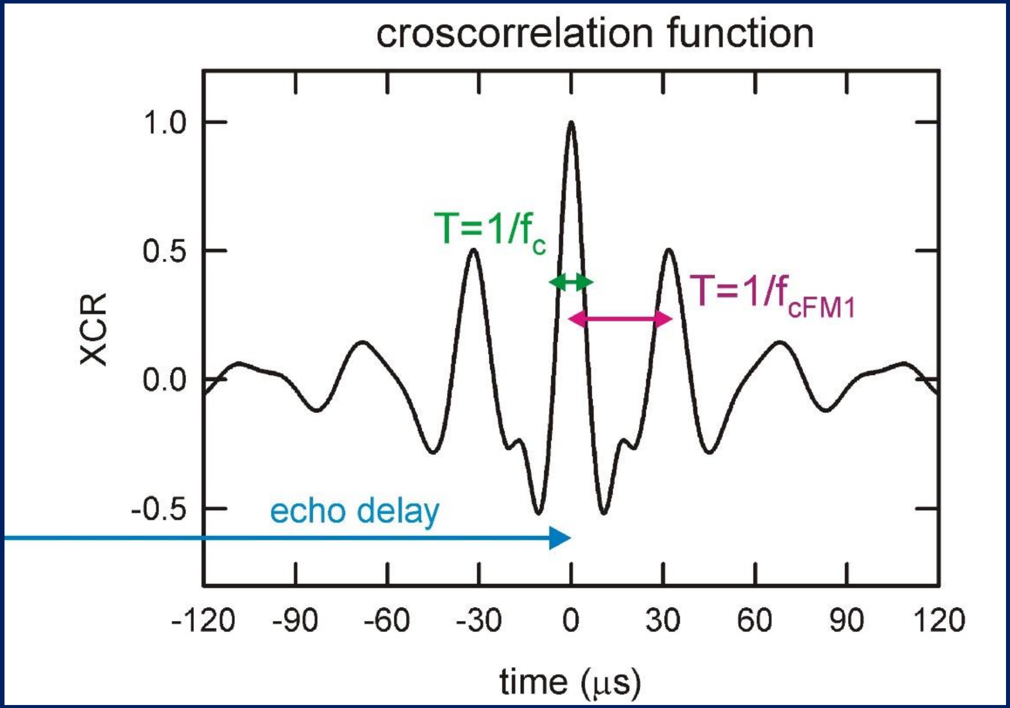


Spectrogram showing representative echoes from the scene. Notice that only the echo from the insect contains the full bandwidth of the broadcast. All other echoes are dominated by lowpass filtering.

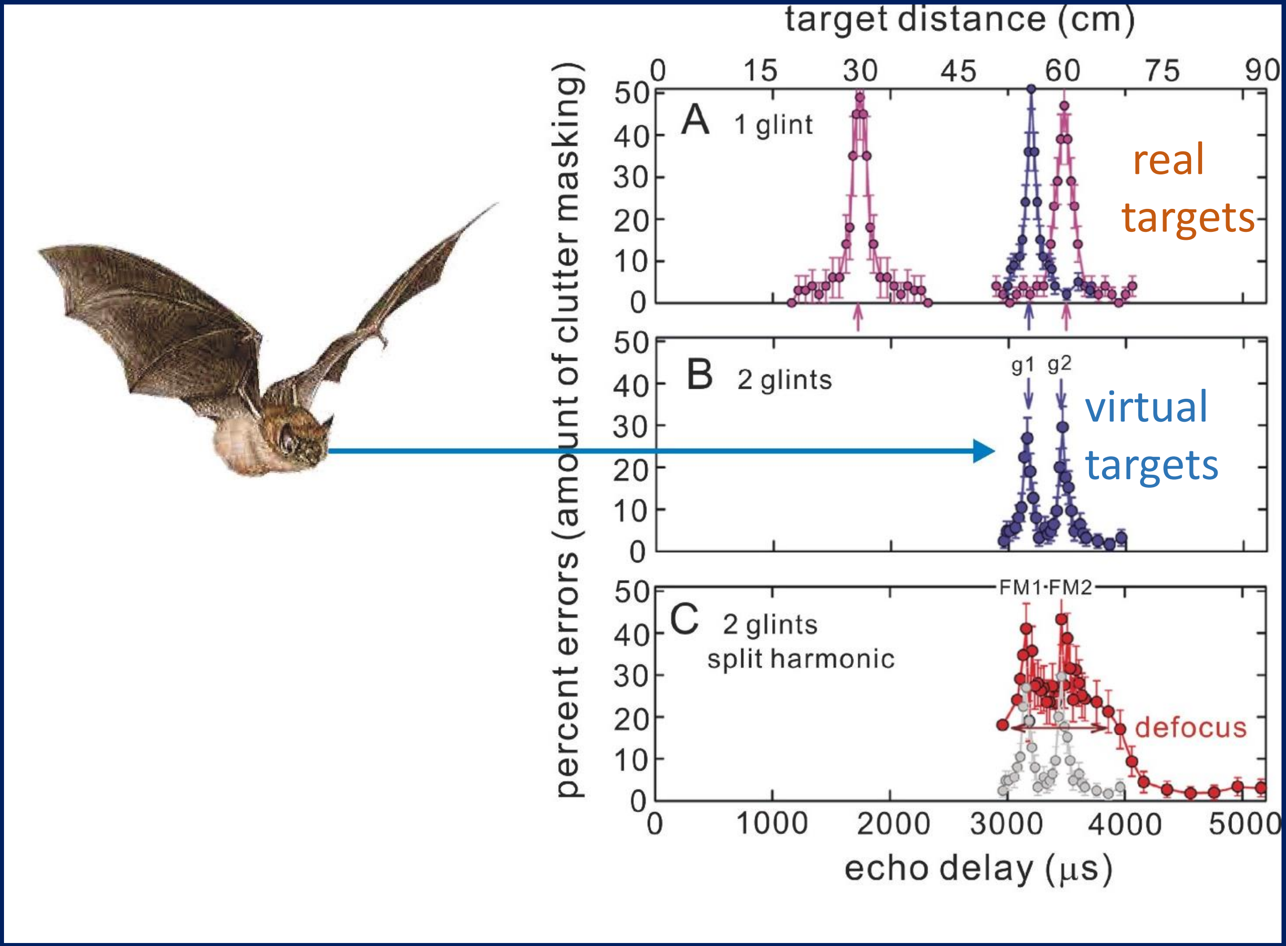
One further factor that affects the spectrum of echoes is the external ear. As is true for most mammals, reception by the external ear imposes an elevation-dependent notch in the spectrum of echoes. In big brown bats, the pinna-tragus notch moves from low to high frequency as elevation increases. It is confined to frequencies in the 1st harmonic, FM1.



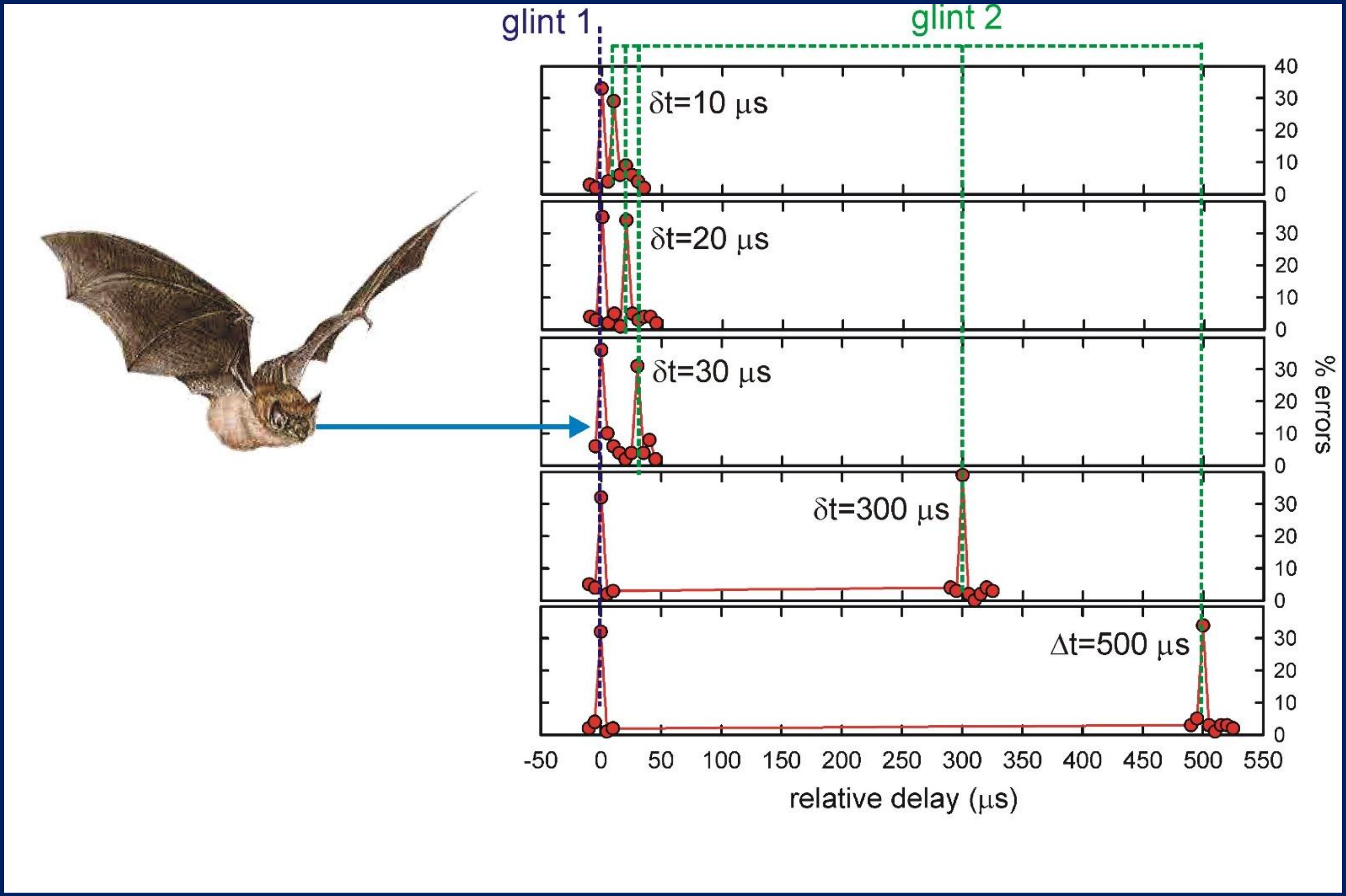




Wideband FM broadcasts are used by radar and sonar systems to determine target range from echo delay very accurately, and to take advantage of the wide span of frequencies to "paint" images that depict target shape. However, the spread of frequencies across the duration of the sweeps has to be counteracted by compressing each pulse-echo pair signal into its shortest form—the crosscorrelation function. This is roughly like wringing out of the signals all of the information not relevant to determining delay.

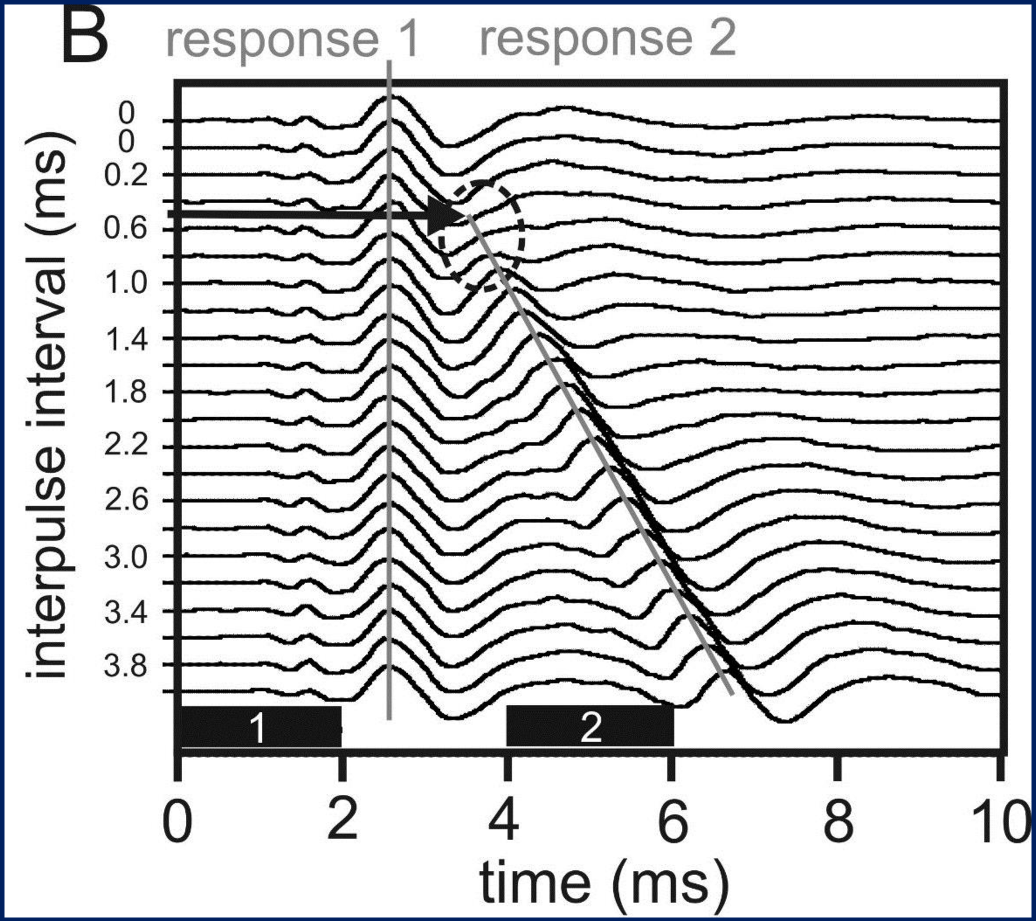


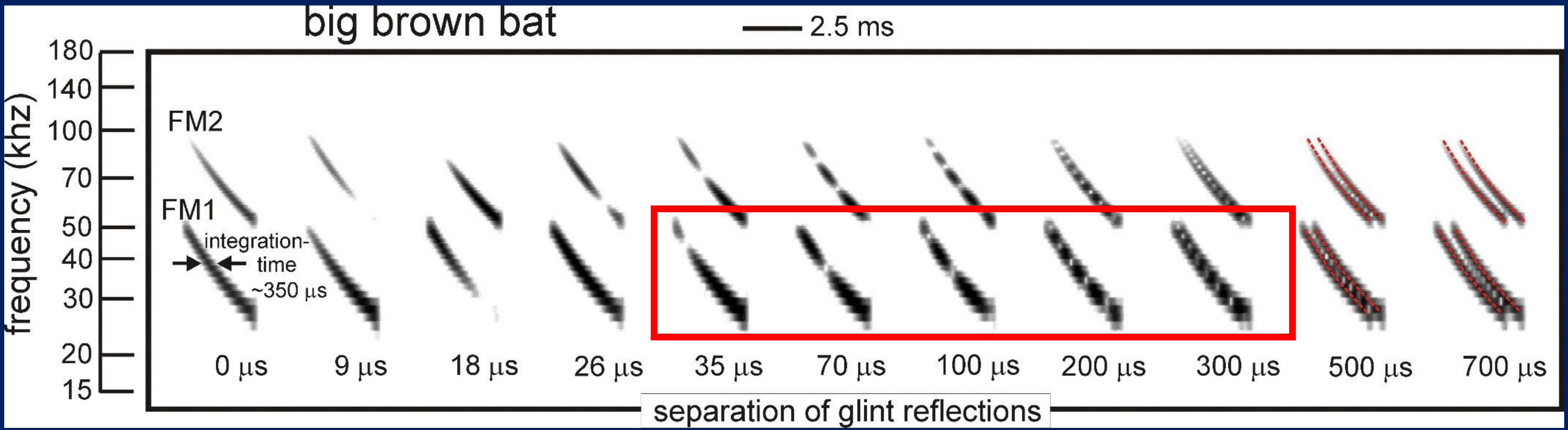
Bats do perceive target range from echo delay. These curves show the percentage errors made by big brown bats in several different range or delay discrimination psychophysical tests. Electronic manipulation of virtual target range confirms that echo delay is used to perceive range.



They also perceive the distance to the different reflecting part of targets, called *glints*. This is the sonar equivalent of the target's shape, expressed here as two-point images.

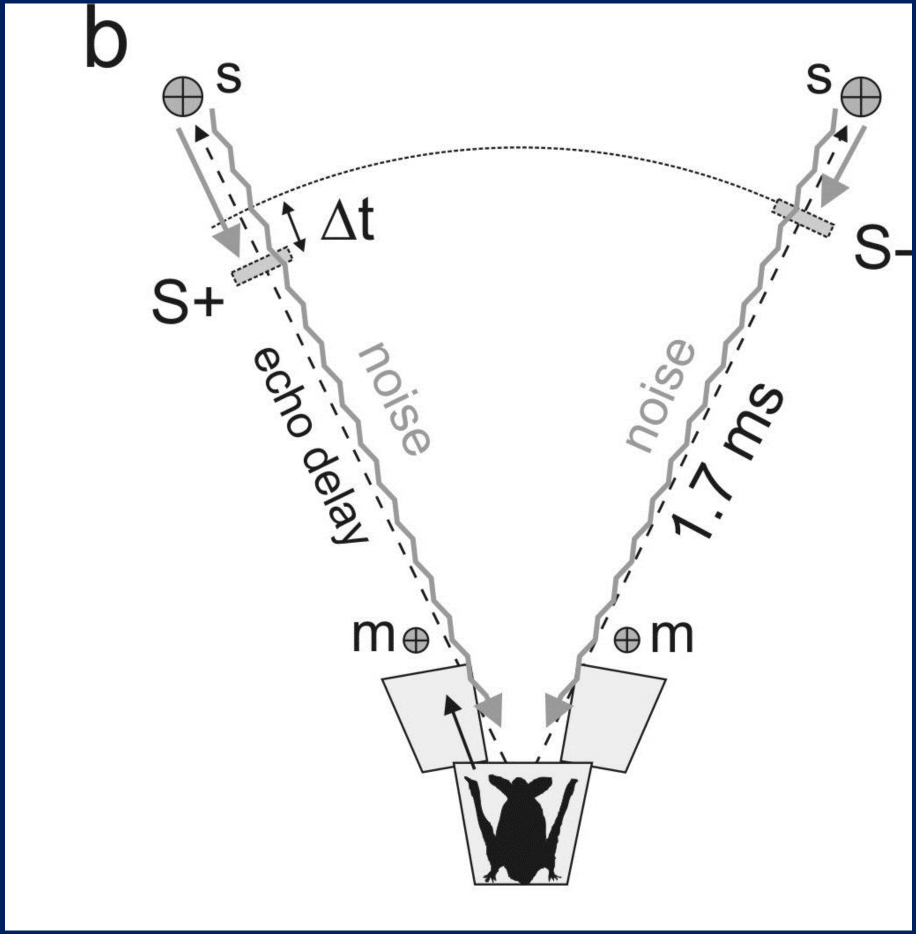
One of the earliest concerns about the neurophysiology of echolocation is recovery time—the duration of time following reception of a sound when neural responses are not yet able to register a second sound. It is relevant for understanding how the delay of the second target glint is registered—by time or by spectrum.

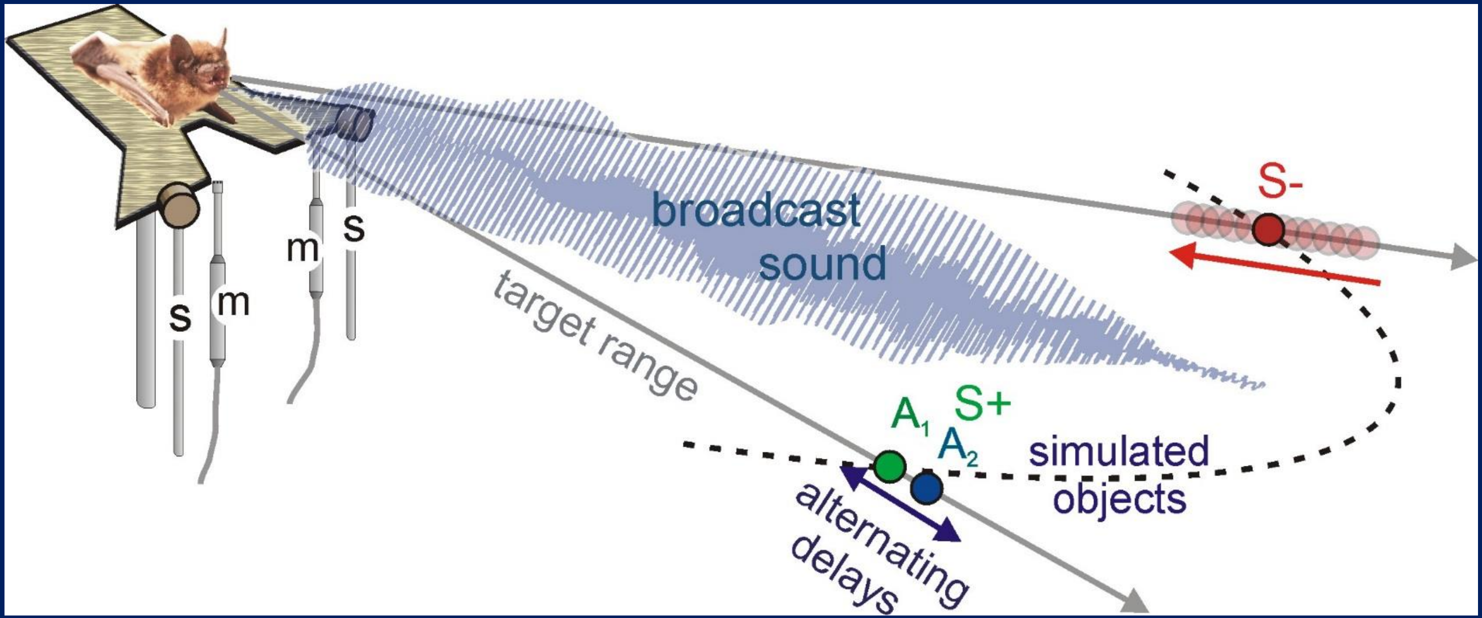




The acoustic cues for resolving the second of two reflections (from a slightly more distant part of a target) varies according to the delay separation of the glint reflections in relation to the integration-time of the bat's auditory spectrograms.

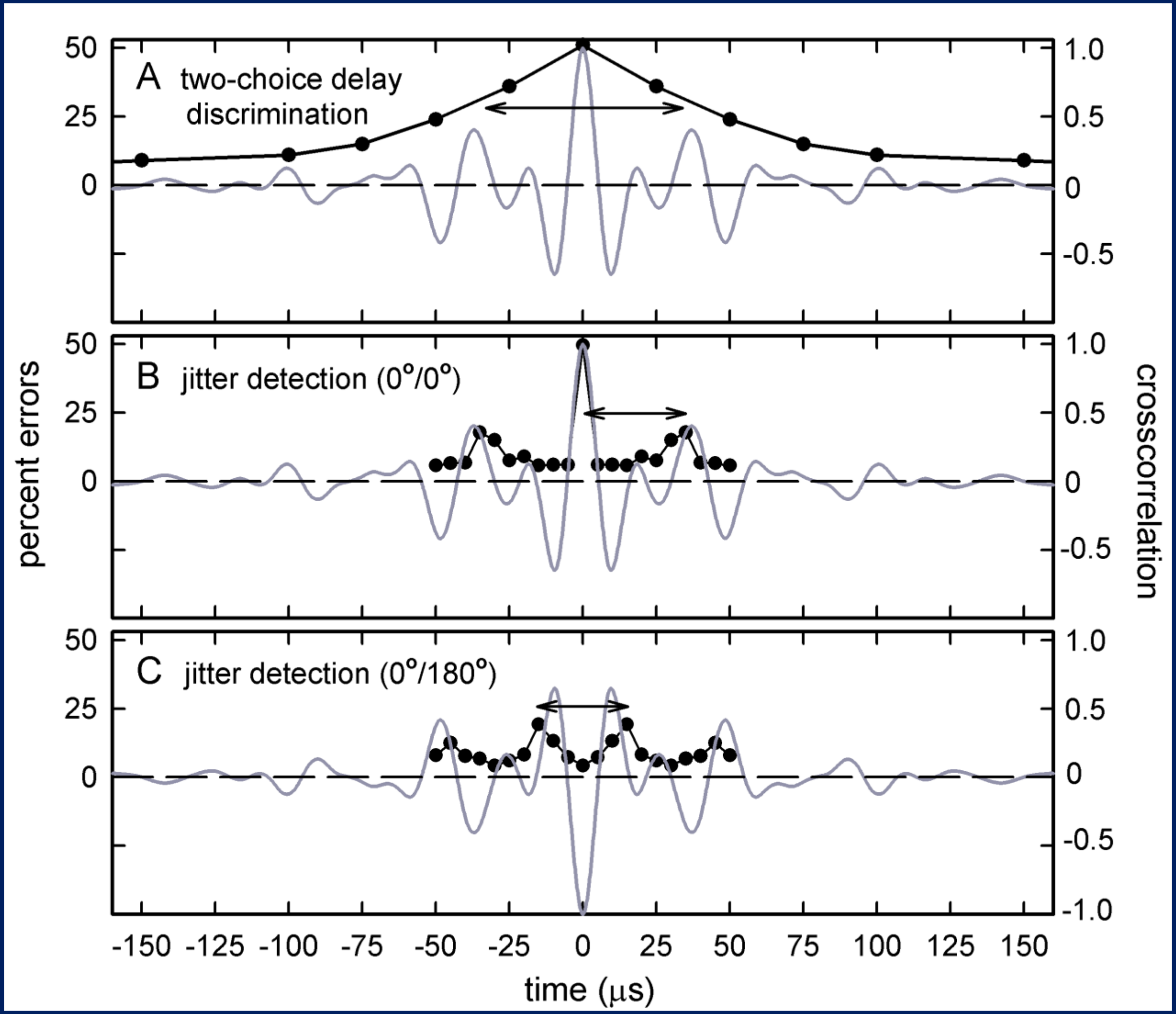
The target range or echo delay discrimination procedure involves the bat moving its head left and right to scan the real or virtual targets so it can choose which is at the correct (rewarded) distance. This induces unwanted perturbations in the distance to the targets, which blurs the performance curve.

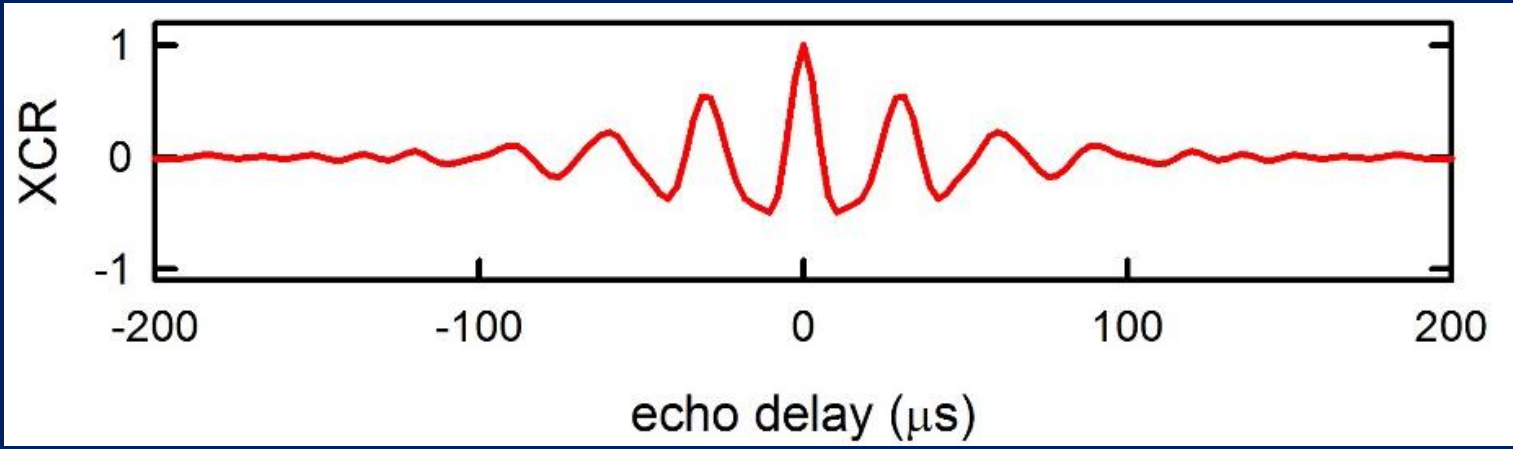


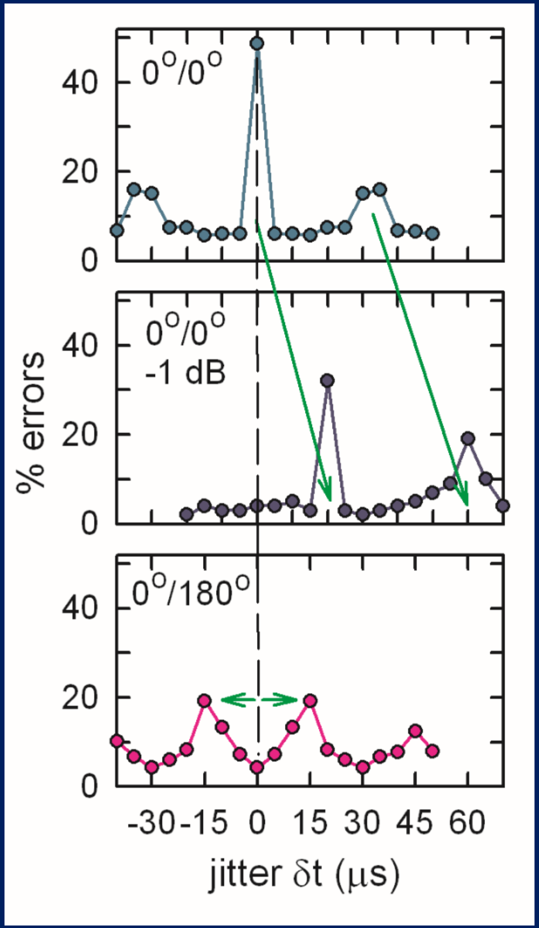


The solution is to alternate, or jitter, the delay of echoes from one broadcast to the next, so the bat’s head remains relatively stationary while the crucial stimulus delay difference is acquired.

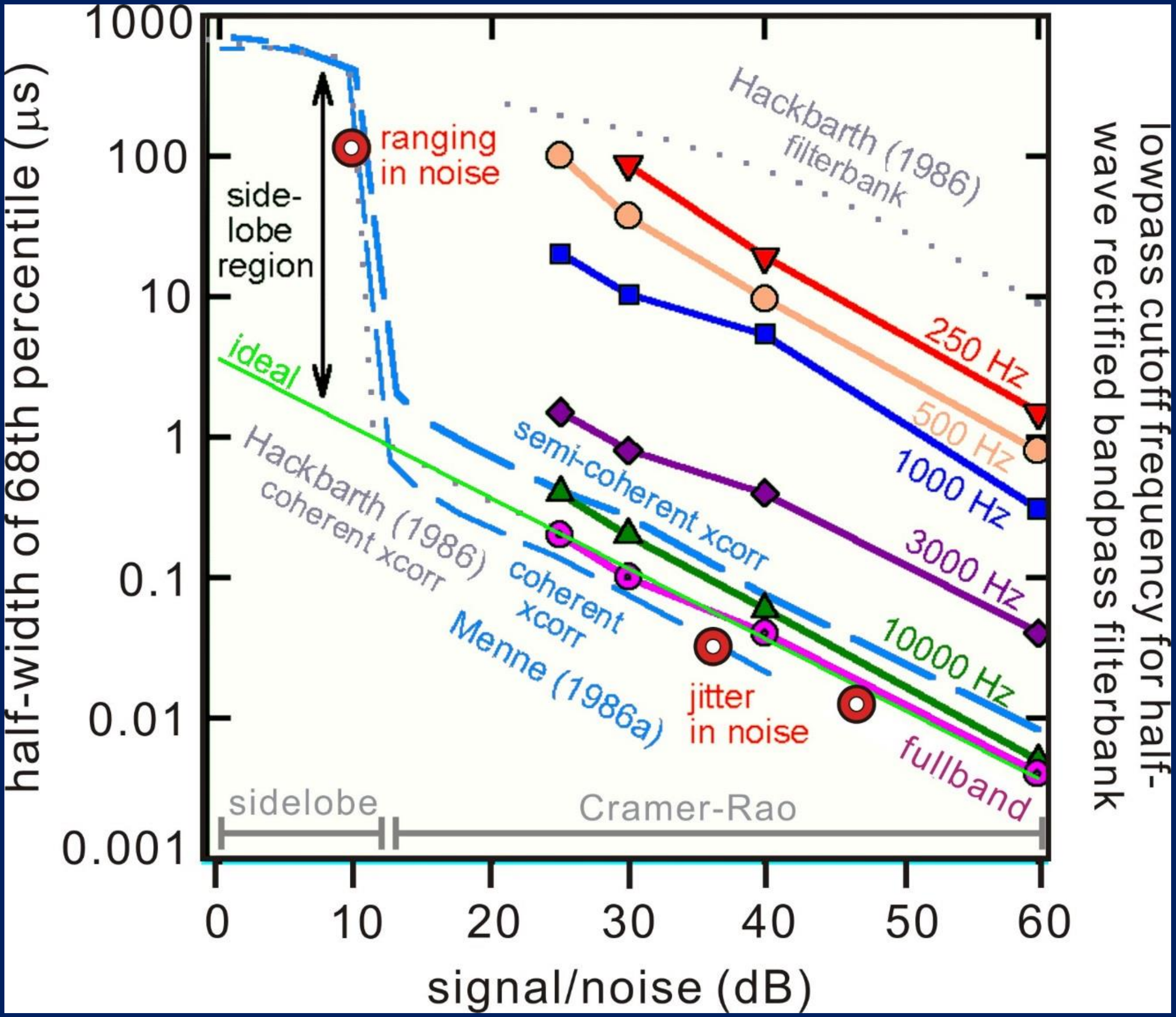
The blurred performance curve from 2 choice delay discrimination tests (A) is replaced by finer delay acuity that traces a curve resembling the half-wave-rectifice crosscorrelation function of the echoes (B). When the jittering echoes also alternate in phase, the performance curve resembles the negative side of the function.



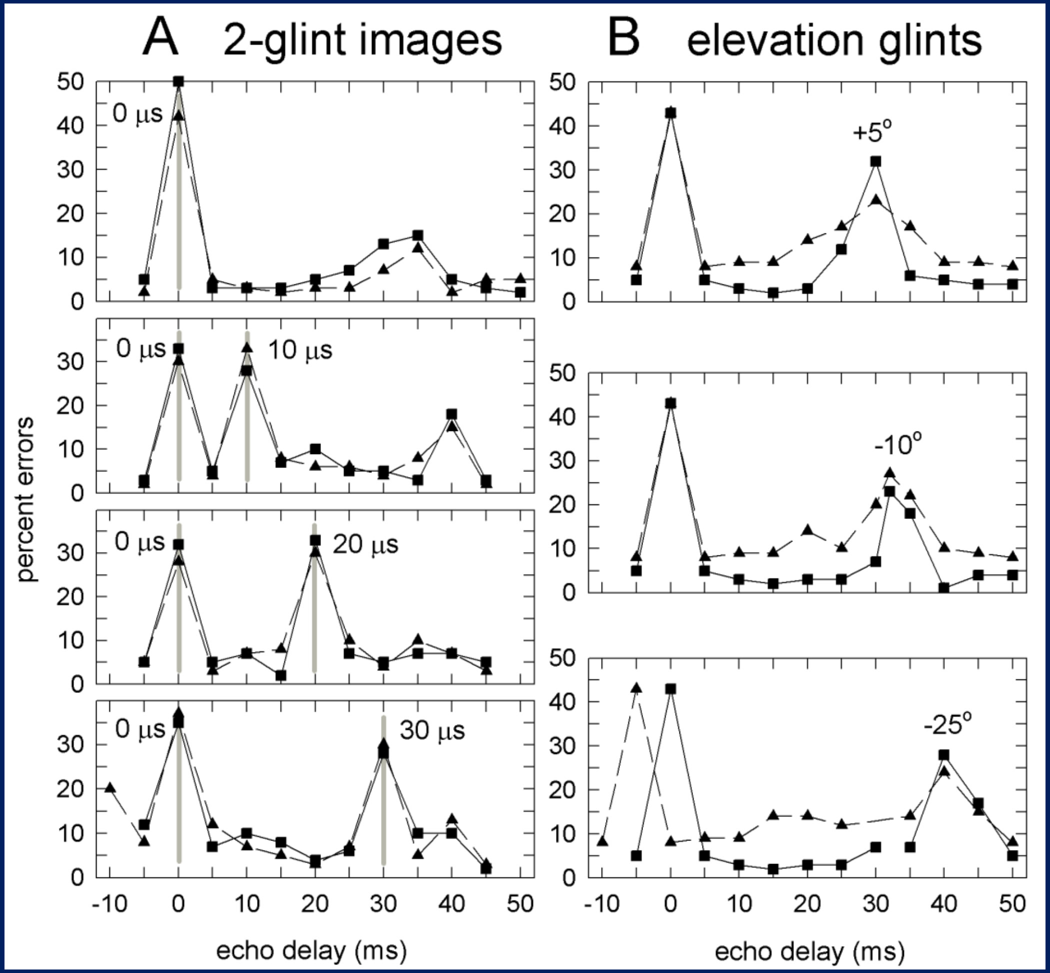




Amplitude-latency trading (@~15 μs/dB) Confirms that the timing of neural responses to echoes conveys echo delay information up the auditory pathway.

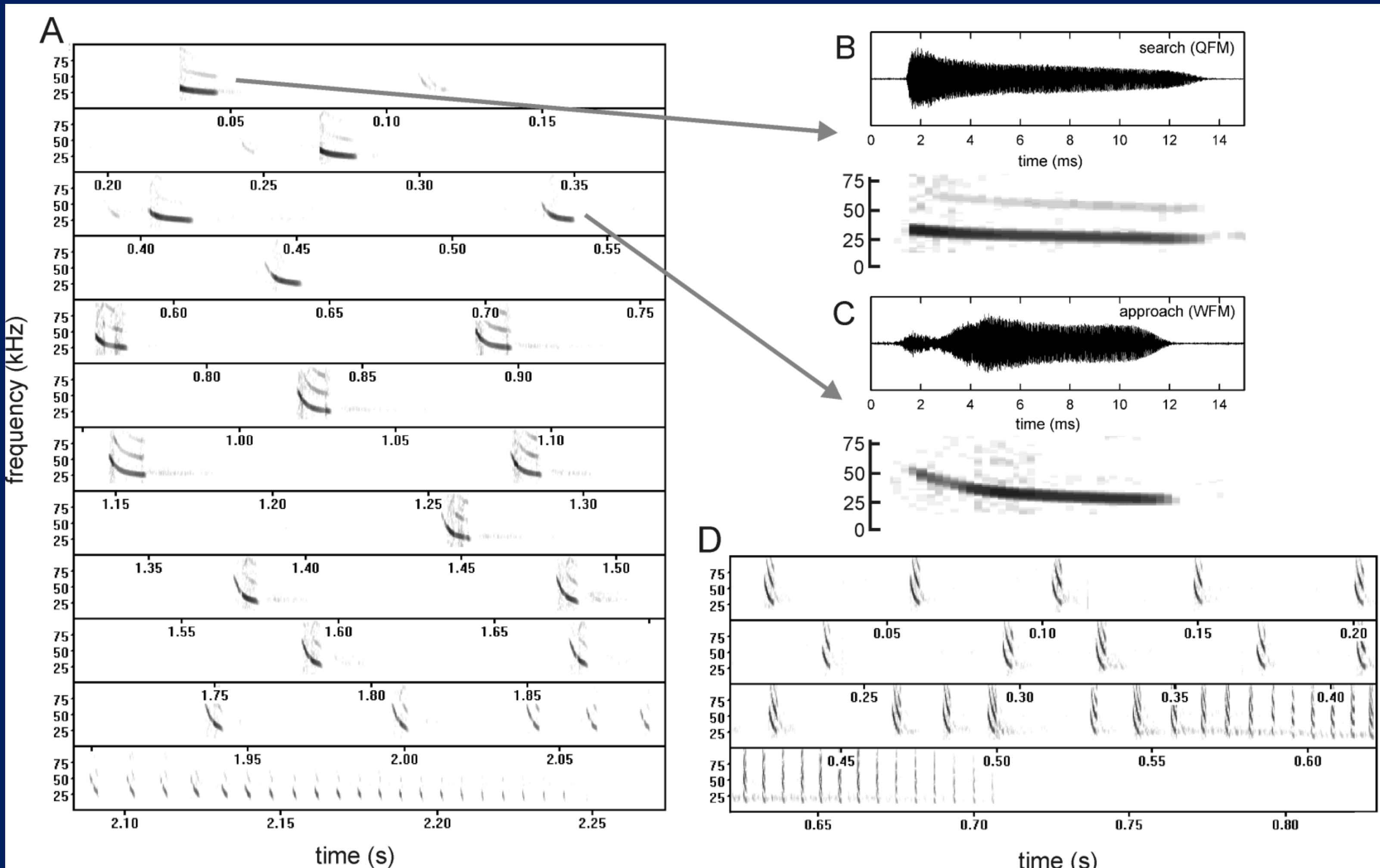


Echo delay accuracy at SNR > 15 dB corresponds to ideal performance (Cramér-Rao bound), but it is based on time-frequency methods, not crosscorrelation. Large-scale statistical simulations reveal that coherent processing is achievable even with 10-12 kHz lowpass filtering of auditory time-frequency representations.



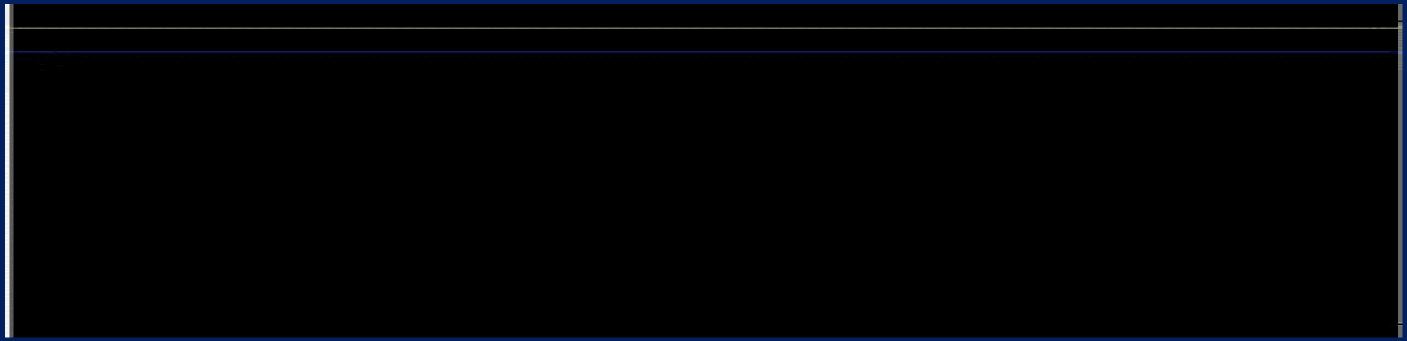
Bats perceive the arrival-time of reflections from different parts of the target (glints). They even perceive the reverberation built into the external ear as a new “glint” that represents target elevation.

Big brown bats are insectivorous and frequently capture prey in aerial interception maneuvers guided by sonar. During interceptions the bat continuously changes the duration, frequency sweep structure, and interpulse interval of its broadcasts.



The continuous nature of the bat's adaptive changes in broadcasts imposes severe constraints on the auditory processing of pulses and echoes. Producing the crosscorrelation function requires a stored template of the broadcast, but if the transmitted waveforms change, the template has to change, too.

The bat's FM broadcasts contain harmonics throughout interceptions.



Aerial capture

capture from vegetation

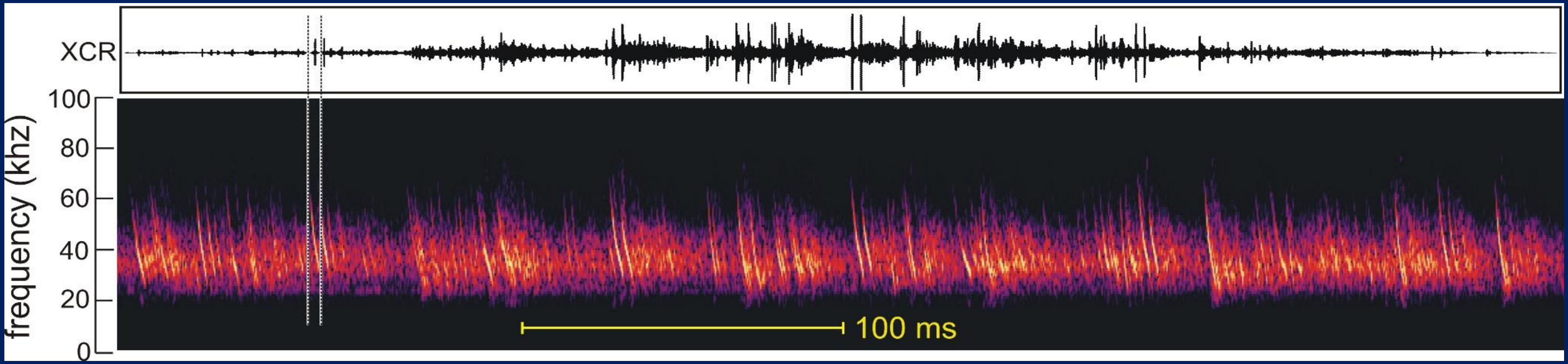
bat-to-bat “dogfight”

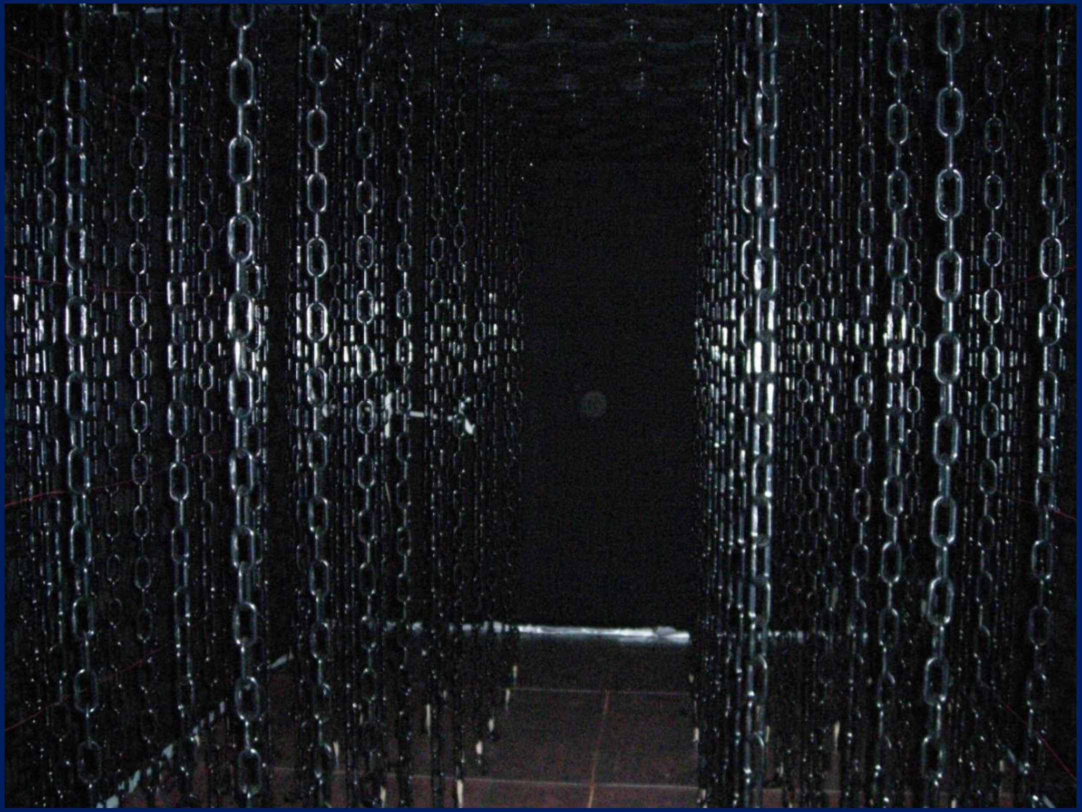
multiple-bat chases

swarming over water

swarming around roost

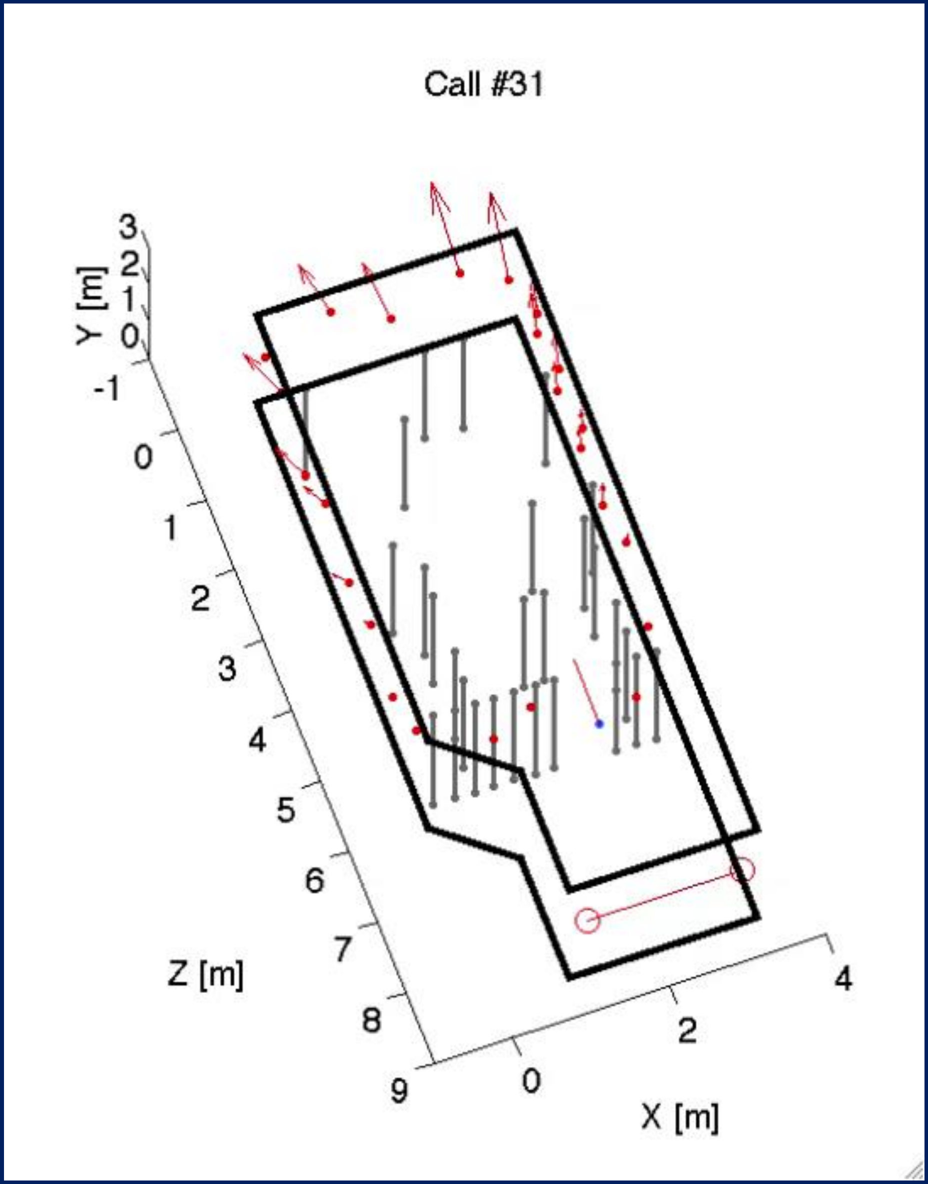
Multiple bats flying in reverberant room.

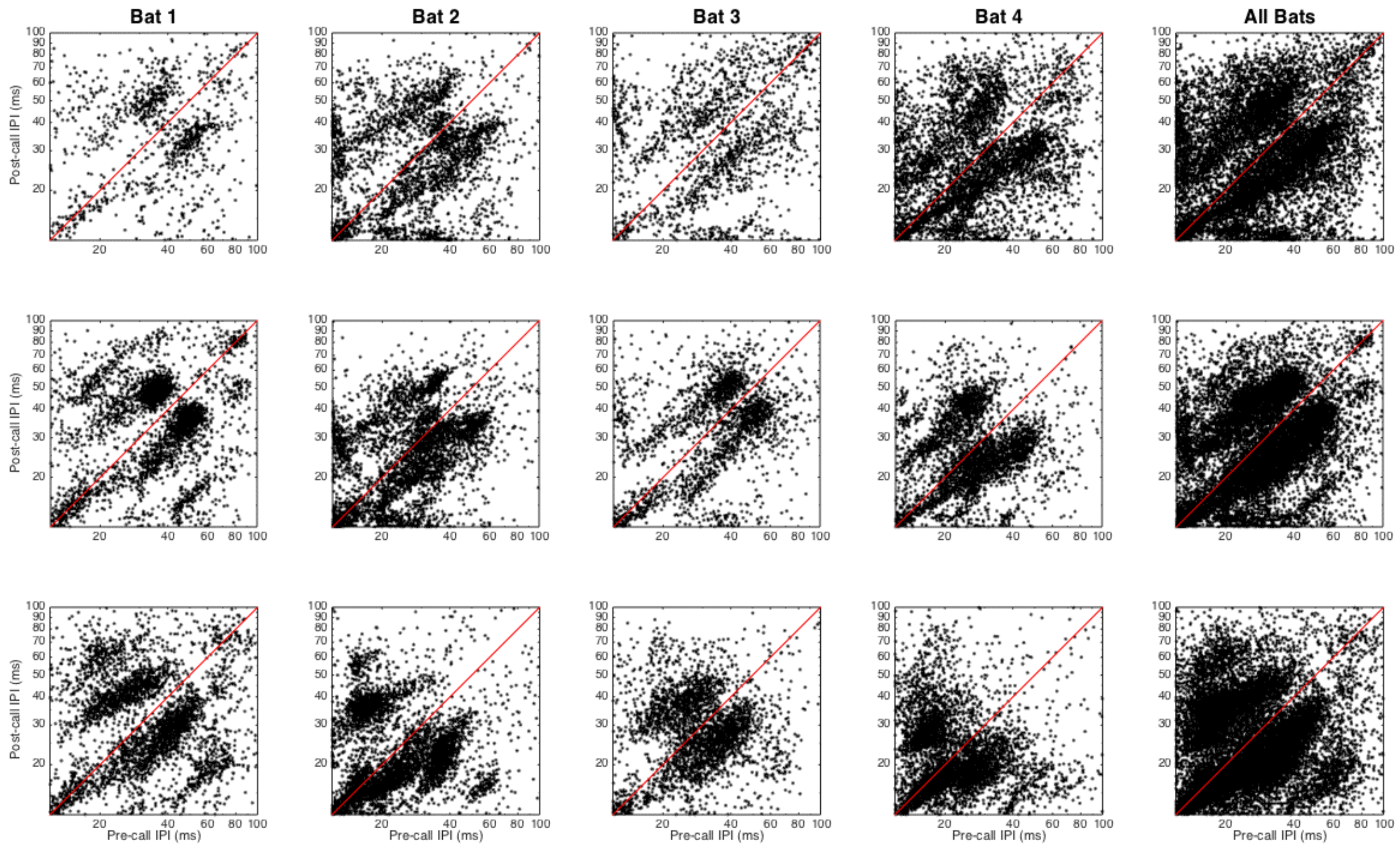




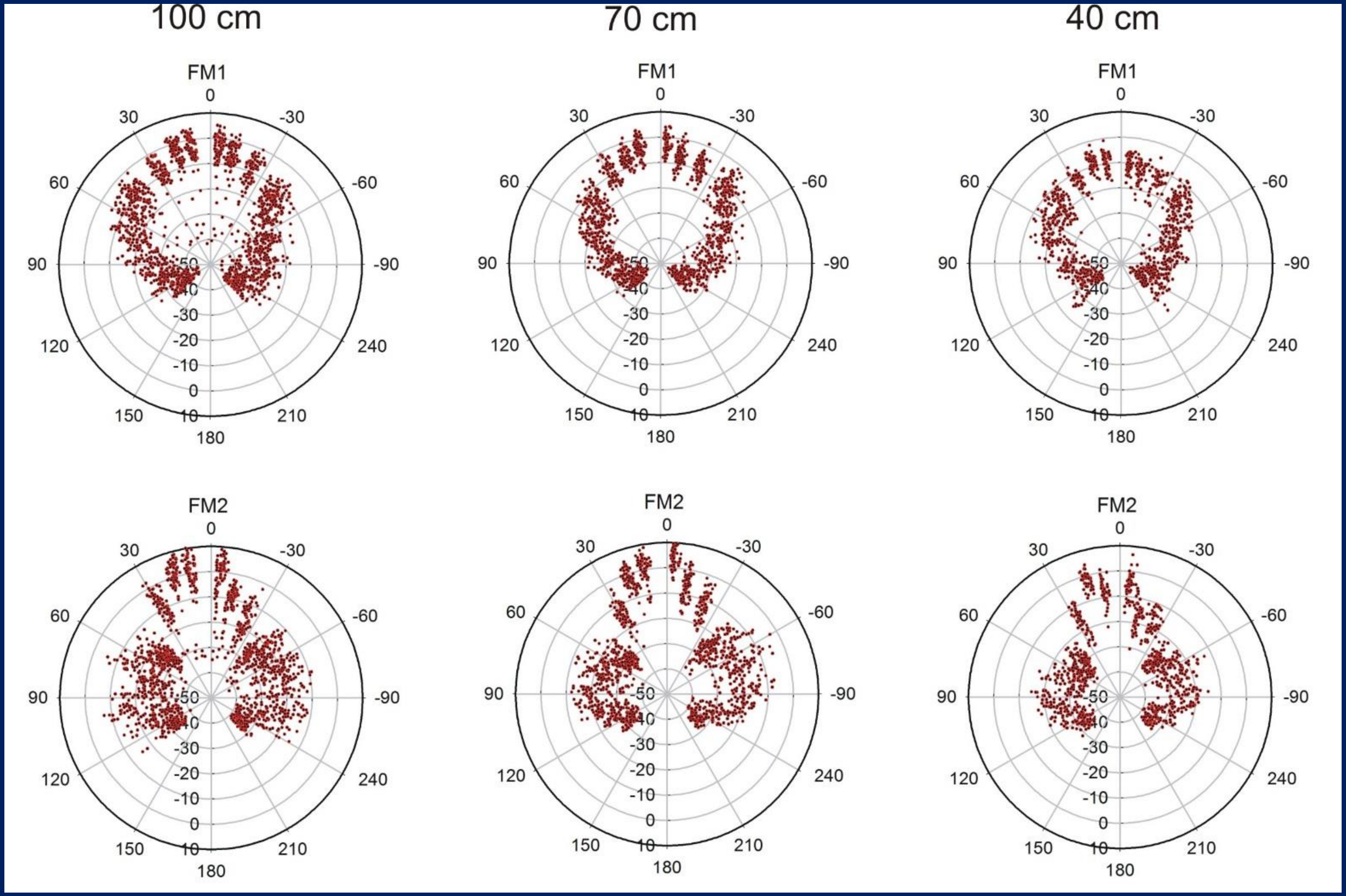
To investigate how bats cope with dense, range-extended clutter such as broad swathes of vegetation, we use obstacle arrays made of different configurations of black plastic chains hanging from the ceiling of a flight room.

In common with several labs, we have used microphone arrays to track bats flying in the array of chains. The microphones trace the flight path, the aim of the sound beam, and the approximate width of the transmitted beam.

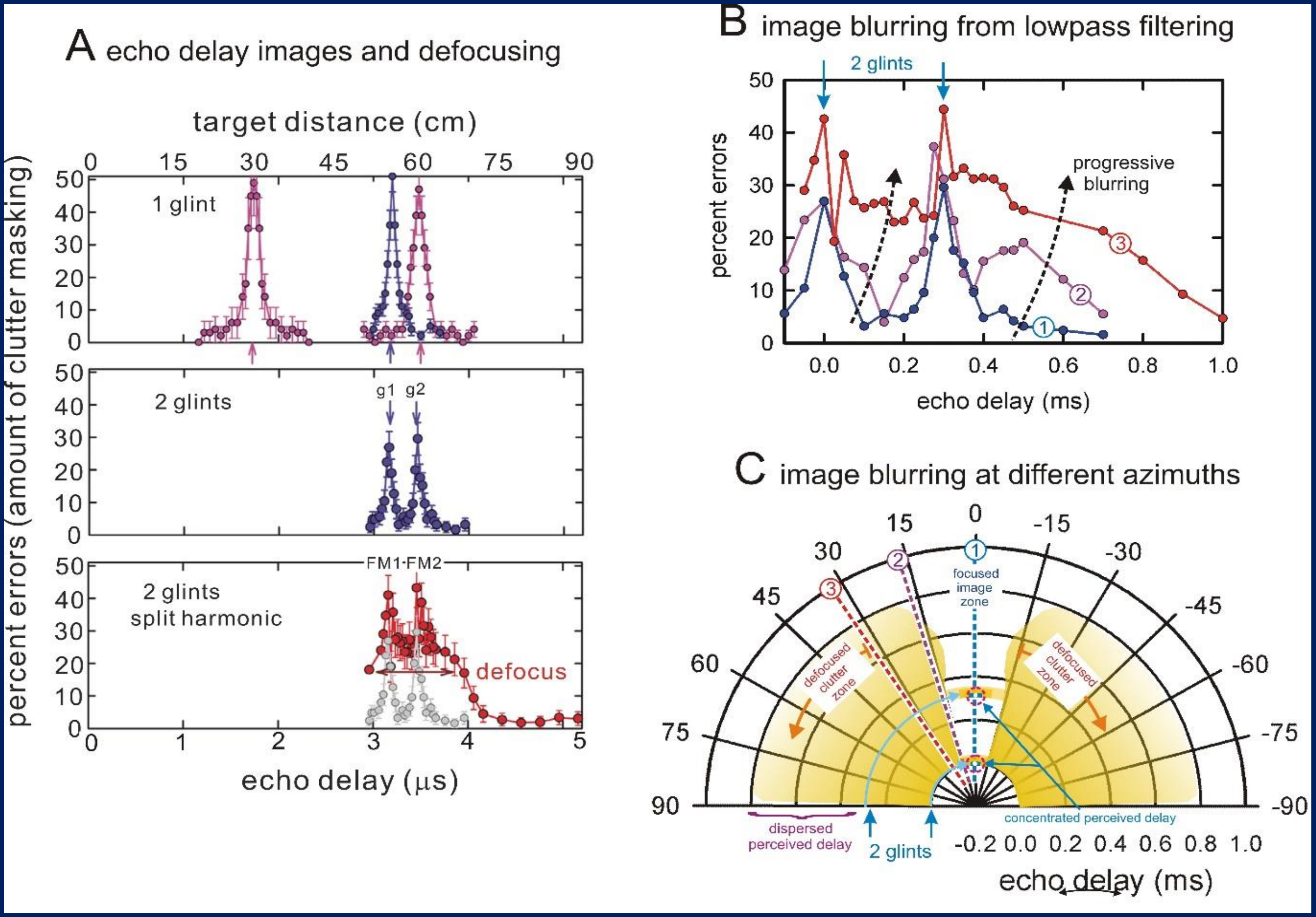




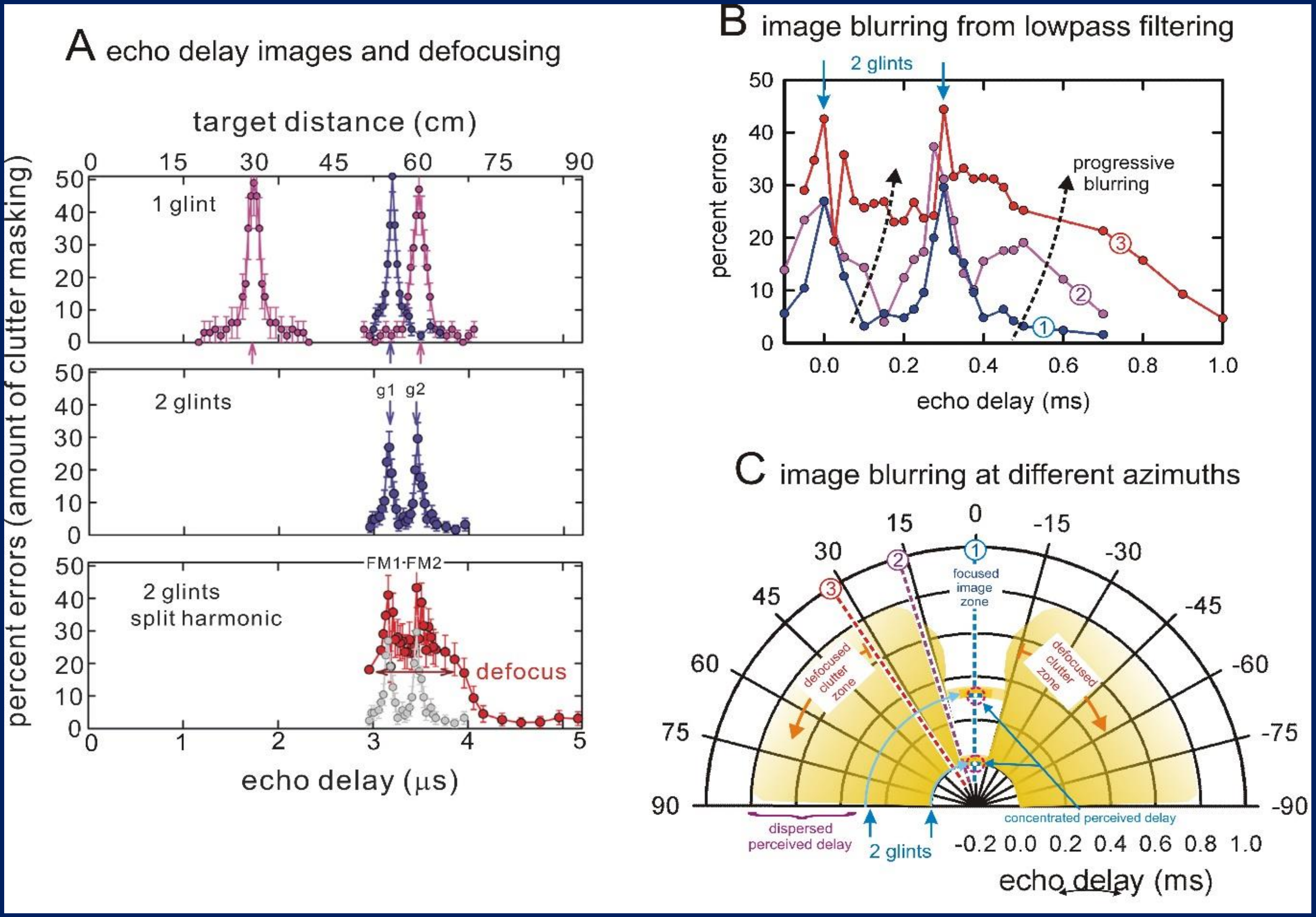
Interpulse intervals alternate between long and short in rougly fixed proportions, with somewhat different patterns in different individual bats.



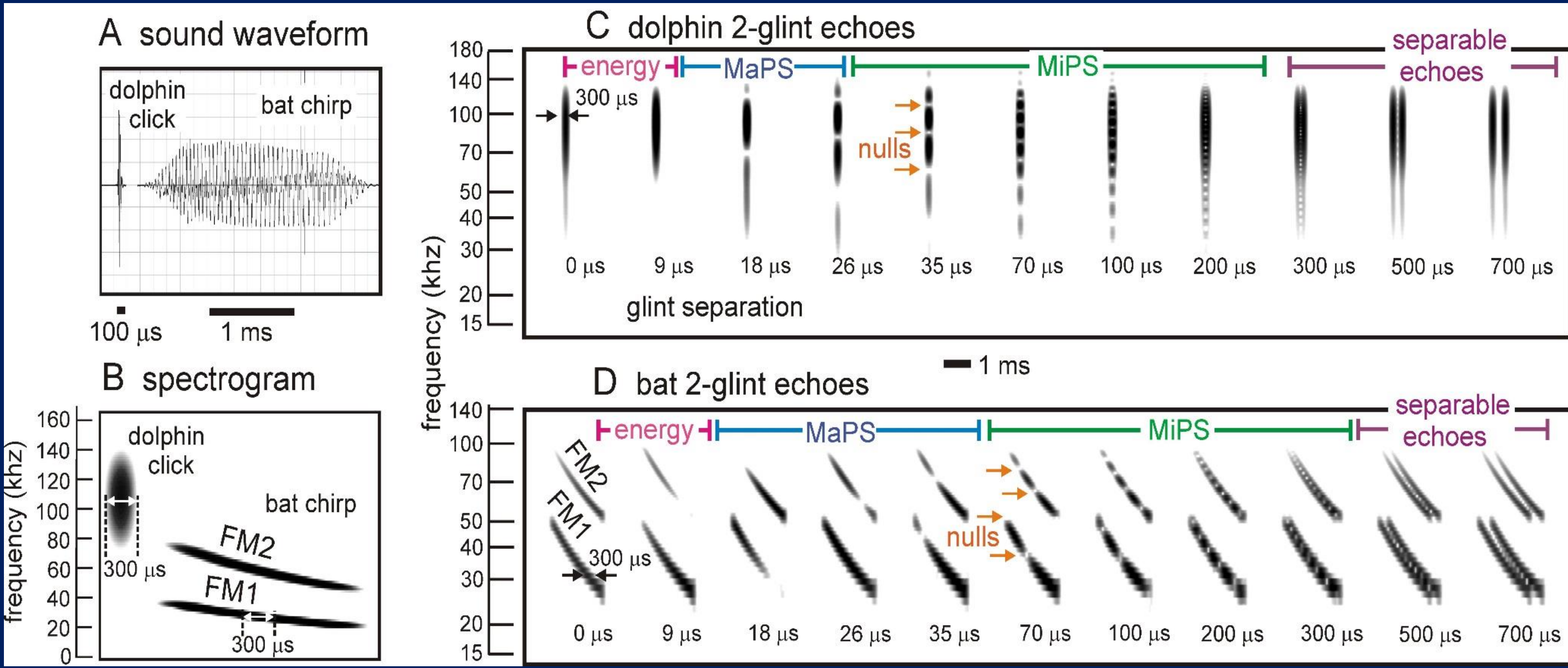
Surprisingly, brown bats do not narrow their beams when the corridor through the clutter becomes more restricted. Instead, they resort to internal beamforming computations on their perceived images.



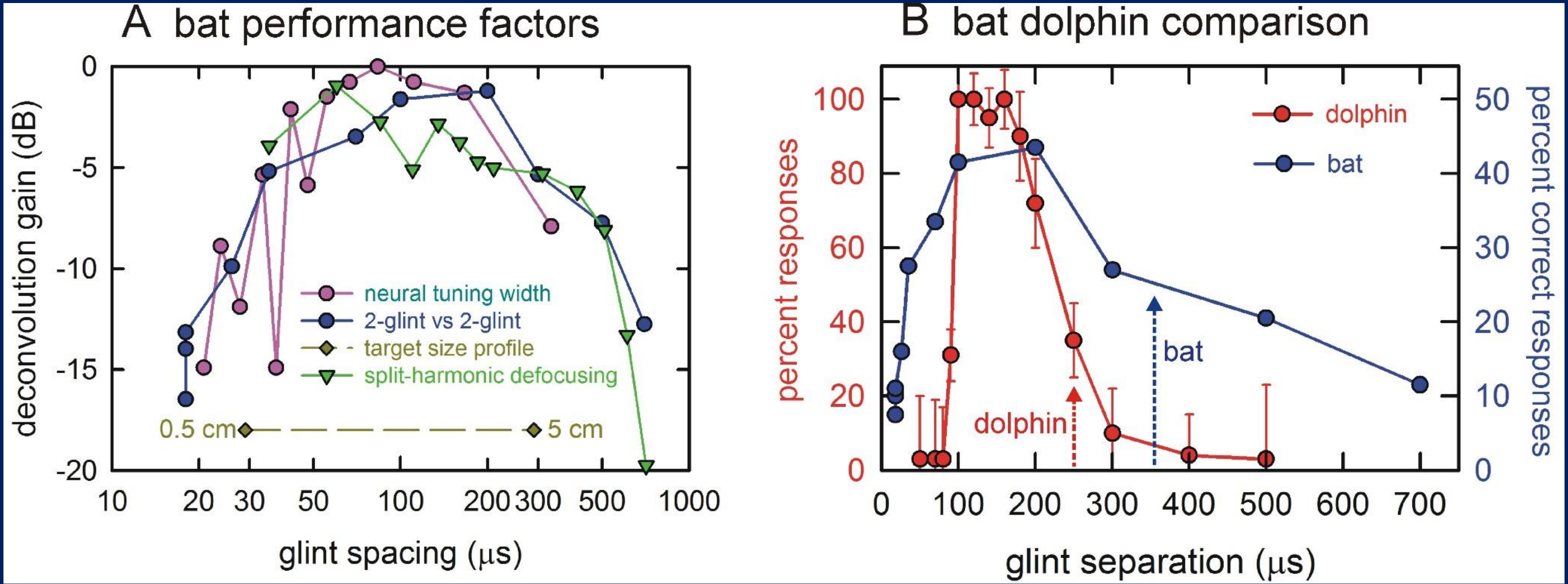
Wideband FM biosonar images are assembled on a scaffold of echo delay estimates. (A) Psychophysical measurements of target range or echo delay images for 1-glint targets, a 2-glint target, and 2-glint echoes with temporal dislocation of the harmonics FM1, FM2). (B) Image defocusing occurs for any disruptions of the neural representation of the harmonics.



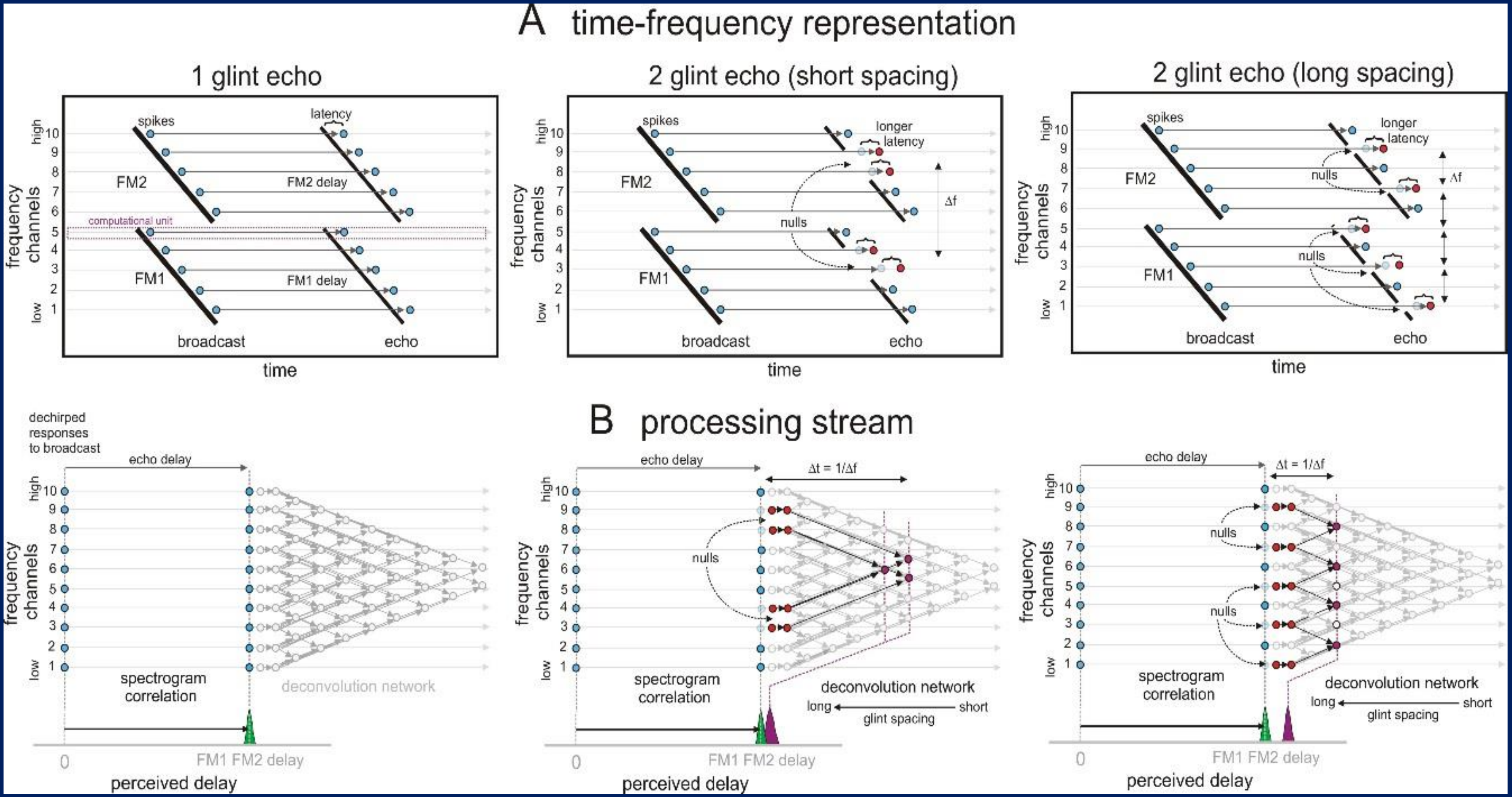
(C) The prevalence of lowpass filtering for echoes from offside or distant clutter causes the neural representation of the harmonics to misalign in time, which defocuses the images of clutter and renders them not capable of interfering with perception of full-band echoes arriving from straight ahead. Objects located on the beam axis are rendered into focused images.



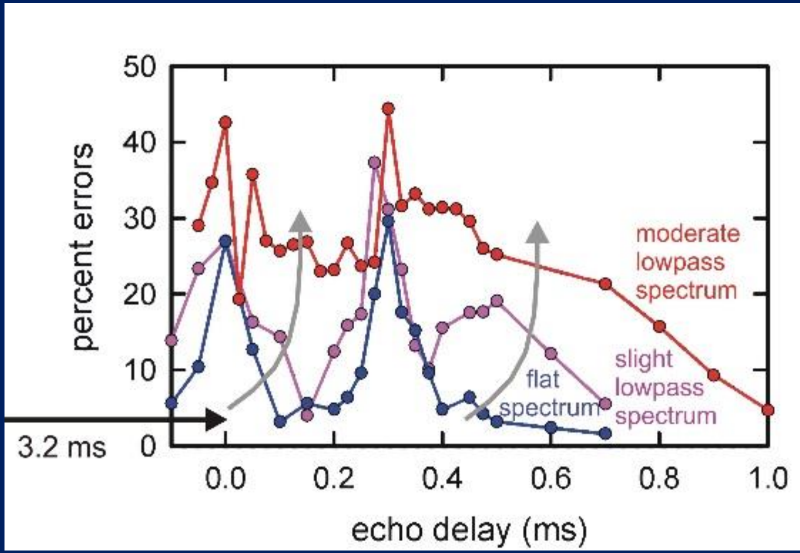
Echo delay resolution involves the interaction between peripheral auditory integration-time and the size of the time interval between overlapping reflections. Big brown bats can resolve two overlapping reflections from about 3 μs to over 700 μs, using a constellation of cues.

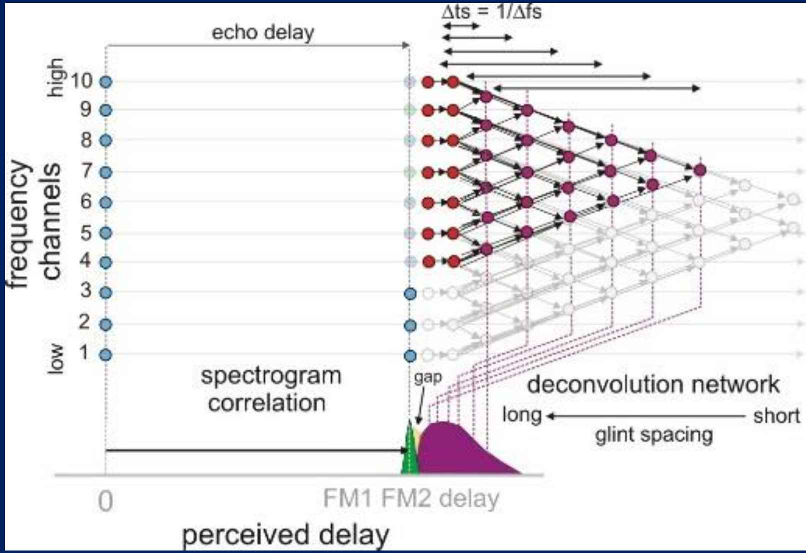


A bat performance factors
0
-5
-10
-15
-20
deconvolution gain (dB)
neural tuning width
2-glint vs 2-glint
target size profile
split-harmonic defocusing
0.5 cm
5 cm
10
20
30
50
100
200
300
500
1000
glint spacing (μs)
B bat dolphin comparison
percent responses
100
80
60
40
20
0
dolphin
bat
50
40
30
20
10
0
percent correct responses
dolphin
bat
0
100
200
300
400
500
600
700
glint separation (μs)

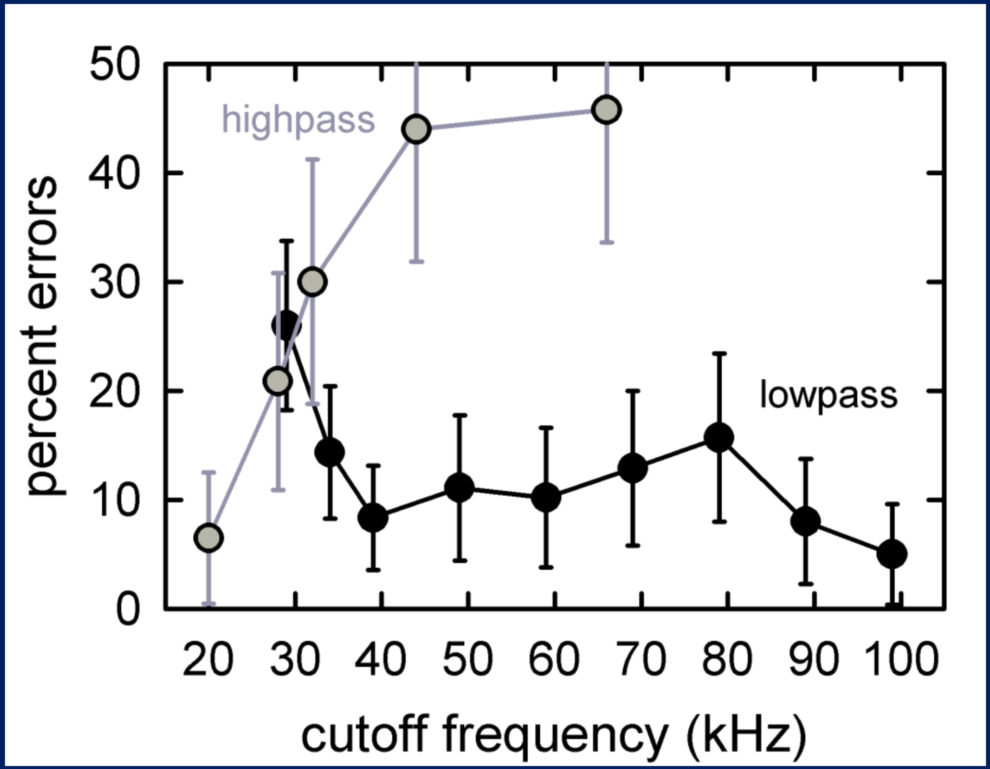


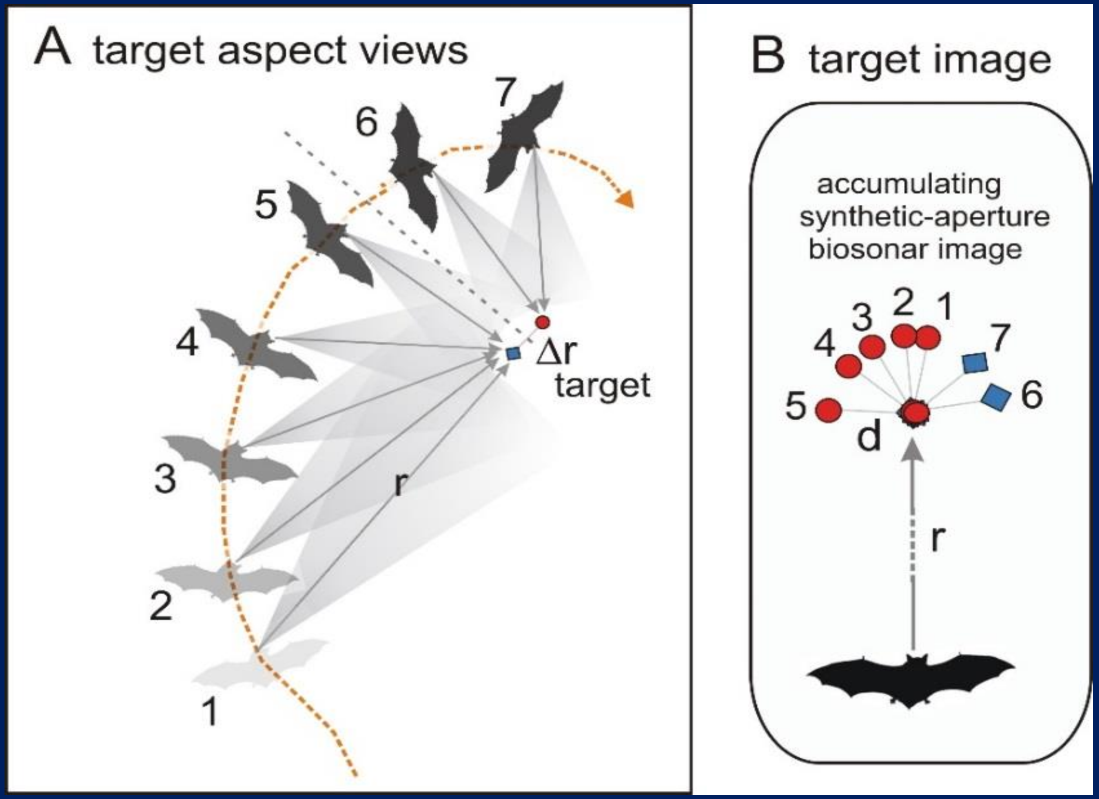
Echo delay images for a flat-spectrum 1-glint echo and two different 2-glint echoes. The overlap and interference between glint reflections places nulls in the echo spectrum at frequency intervals reciprocal to the glint separations. The longer latencies of responses at null frequencies activate the triangular deconvolution network, which determines the glint separation and writes it into the images.



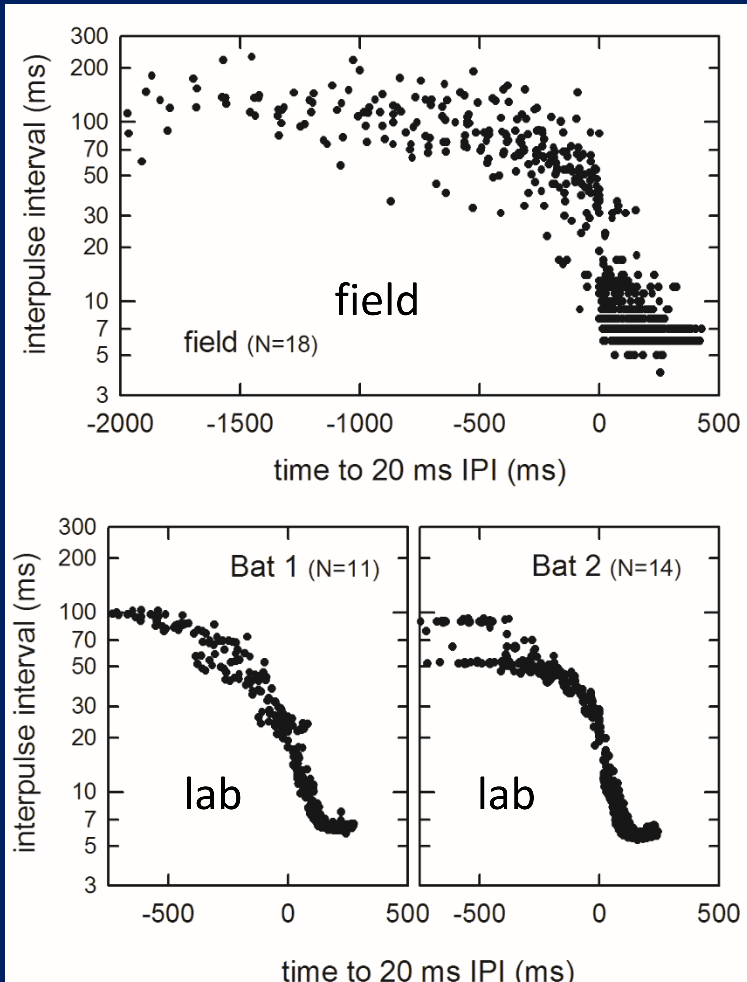


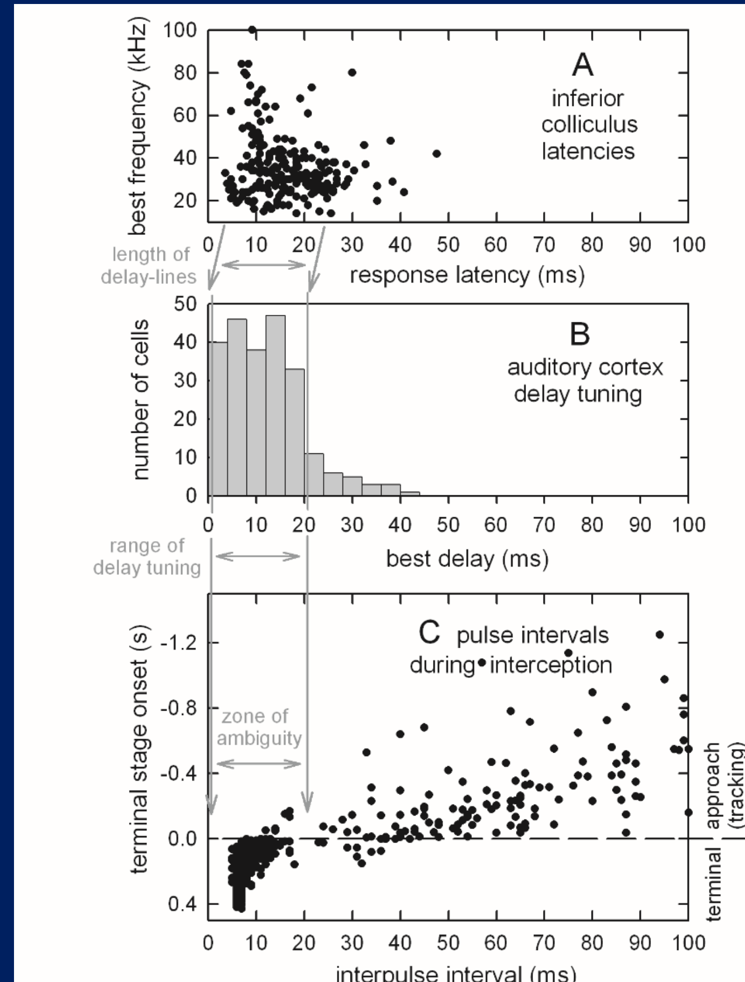
Lowpass filtering overwhelms the deconvolution network by initiating the registration of many spectral nulls and their corresponding glints. This defocuses the images derived from lowpass filtered echoes.


highpass
lowpass
percent errors
0
10
20
30
40
50
20
30
40
50
60
70
80
90
100
cutoff frequency (kHz)



Inclusion of the target's glints in the delay images allows a kind of synthetic aperture reconstruction of the target's shape from successive "looks" at the object being viewed along a curving flight path. The reconstruction is built up from the nearest glint, which shifts with aspect angle without disrupting the reconstruction process.





Field and labs sound emission patterns differ in the trajectory of interpulse intervals during interceptions. In open-area field maneuvers, IPIs skip over the range of 16-30 ms and move directly to the terminal stage. The terminal burst of sounds at shortest IPIs appears to ignore pulse-echo ambiguity, and the relation to neural responses suggests the burst is treated as a single, extended broadcast.