



Auditory Localization of Ground-Borne Vibrations in Snakes

Paul Friedel,¹ Bruce A. Young,² and J. Leo van Hemmen¹

¹*Physik Department, Technische Universität München, 85747 Garching bei München, Germany*

²*Department of Biology, Washburn University, Topeka, Kansas 66621, USA*

(Received 19 July 2007; revised manuscript received 19 November 2007; published 28 January 2008)

Interaural time differences allow many animals to perform azimuthal sound localization. Snakes lack a tympanic membrane, external ear openings, and any other superficial indication of an auditory mechanism. They do, however, possess an inner ear with functional cochlea. The oval window is connected through a loss-free osseous lever system to the two, *de facto* independent, sides of the lower jaw, which typically rest on the substrate. The footfall of prey generates small-amplitude, low propagation velocity, Rayleigh waves in the soil. This type of wave can be described as fluid motion. Accordingly we apply naval-engineering techniques to show that lower-jaw motion gives rise to a neuronal representation of the auditory world with realistic sensitivity and stereo precision.

DOI: [10.1103/PhysRevLett.100.048701](https://doi.org/10.1103/PhysRevLett.100.048701)

PACS numbers: 43.80.+p, 43.60.+d, 68.35.Iv, 84.35.+i

The ability of snakes to hear has been debated for decades, even though Hartline [1] and Wever [2] have shown that acoustic stimuli produce responses in the cochlea and brain of snakes. Those who rejected hearing in snakes noted that snakes lack an outer ear, the absence of which was assumed to hamper acoustic reception, and that reliable descriptions of snakes' behavioral responses to sound were scarce [3]. Part of the debate has focused on the question of whether snakes are able to receive airborne, or only ground-borne, stimuli. Evidence for behavioral responses to both airborne and ground-borne vibrations has been found recently [3]. In this Letter, we focus on biophysical principles underlying the detection of ground-borne vibrations and neuronal processing of the ensuing auditory signals by snakes.

Young and Morain [4] demonstrated that the horned desert snake *Cerastes cerastes* can localize and strike live mice using only vibrational cues. The footsteps of a mouse cause surface waves to propagate in the sand, which are then detected by the snake. Figures 1 and 2 illustrate a possible anatomical route responsible for this form of auditory sensitivity. We hypothesize that surface vibrations are picked up by the two sides of the snake's lower jaw. Because of the mechanical connection from the lower jaw to the endolymph of the inner ear, via the linked *quadrate* and *stapes* bones (Fig. 2), lower-jaw vibration can stimulate the cochlea through a lever construction without impedance matching problems as in airborne hearing. To enable the snake to swallow large prey, the two sides of the lower jaw are only loosely connected and can therefore move independently. Hence, *stereo* response to incoming vibrations and sound source localization using interaural time difference is possible.

Both Hartline [5] and Wever [2] have measured the snake's auditory response to vibrations applied to the lower jaw. The response threshold was found to be of the order of 1 \AA (10^{-10} m) with highest sensitivity for frequencies around 300 Hz. This high sensitivity, despite the

presence of a soft-tissue ligament between quadrate and stapes, might be explained by the "stiffening effect" of high-frequency vibrations [6]. Although Hartline and Wever suggested that the jaw-quadrate-stapes pathway presents an efficient way to excite the cochlea, they dismissed the idea that substrate-based hearing is important. We will show below that surface waves may well be effective in exciting the lower jaw and therefore should be regarded as significant sensory input.

The input to the detection system is provided by prey-generated substrate waves. Measurements have shown that around 70% of the energy caused by small disturbances of the sand surface radiates away in the form of *Rayleigh waves* [7]. The remaining 30% is carried away by volume waves propagating into the sand. Small-amplitude Rayleigh waves behave similarly to water surface waves [8], the sand particles carrying out an elliptic motion that is attenuated exponentially with depth. The wave propagation velocity is low, about 45 m/s. Because of strong attenuation at high and low frequencies, the frequency spectrum of Rayleigh waves peaks below 1000 Hz. This

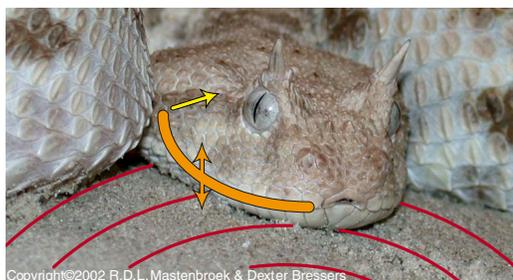


FIG. 1 (color online). The horned desert viper *Cerastes cerastes* typically rests its head on the sand surface to listen for prey. An incoming sand surface wave sets the two independent sides of the lower jaw in motion which is then relayed through quadrate and stapes (Fig. 2) into the inner ear.



FIG. 2. In snakes, two separate (high-contrast inset), very loosely connected sides of the lower jaw are mechanically coupled to the inner ear via the quadrate and stapes bones. Vertical or rocking motion of the lower jaw sides (heave and pitch in Fig. 3) results in effective stimulation of the stapes. Specimen from Zoologische Staatssammlung München.

means that the snake's auditory sensitivity matches the typical surface wave frequency content.

We now estimate the lower jaw's response to incoming surface waves. In so doing, we make three assumptions. First, we treat the substrate as a *continuous* medium. This is justified by the fact that the wavelength of the Rayleigh waves (15 cm) is much larger than the typical sand grain size (250 μm) and the wave amplitude ($\sim 1 \mu\text{m}$) is much smaller than the typical grain size. The clear separation of length scales means we can ignore the movement of individual sand grains [9]. Second, we model the two sides of the lower jaw as two bodies (viz., cylinders) resting on the surface and moving *independently* of each other. We can then apply techniques from naval engineering [10] to calculate the motion response to incoming surface waves. Third, we assume that the Rayleigh wave is not altered much by passing the snake head. This assumption is justified since the wavelength is much larger than the head size of the snake so that bending and refraction effects can be ignored.

Figure 3 shows the geometry of the problem. Jaw motion can be decomposed into six components. Vertical motion (heave and pitch) should be especially effective in eliciting auditory responses (Fig. 2). The equations of motion

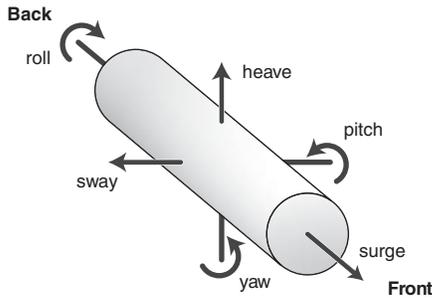


FIG. 3. Movement of a floating body can be decomposed into six components. Motion along three orthogonal axes as well as rotation about these axes. Heave and pitch are effective in stimulating the jaw-quadrate-stapes pathway (Fig. 2).

(EOMs) describing heave u and pitch θ are [10]

$$M\ddot{u} = \int dx f(x, t, u, \theta) \quad M = \int dx m(x), \quad [\text{heave}] \quad (1)$$

$$\left. \begin{aligned} J\ddot{\theta} &= \int dx (x - \bar{x})f(x, t, u, \theta) \\ J &= \int dx (x - \bar{x})^2 m(x). \end{aligned} \right\} [\text{pitch}]. \quad (2)$$

The coordinate x runs along the length of one side of the jaw and \bar{x} is the jaw center of mass (COM). Since the jaw is a *slender* body, an effective one-dimensional description can be used. M and J denote the total jaw mass and moment of inertia, respectively, and $m(x)$ designates the cross-sectional mass density as a function of position x along the jaw. The force f acting on the jaw is given as a function of the difference between sand surface displacement h_{sand} and jaw displacement h_{jaw} ,

$$\begin{aligned} f &= m_H(x)\ddot{h} + n(x)\dot{h} + c(x)\tilde{h}, \\ \tilde{h} &= \tilde{h}(x, t) = h_{\text{sand}} - h_{\text{jaw}}. \end{aligned} \quad (3)$$

The three terms in the expression for the force density f describe inertia, friction, and buoyancy, respectively. The prefactor m_H is called the hydrodynamic mass density, and is determined geometrically from the cross-sectional shape and mass distribution of the floating body under consideration. For a homogeneous cylinder the hydrodynamic mass density equals the ordinary mass density ($m_H = m$) and for a large variety of other cross-sectional shapes one finds $m_H \approx m$.

The surface wave elevation h_{sand} along the jaw is

$$h_{\text{sand}} = a_{\text{sand}} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) = a_{\text{sand}} \cos(kx \cos \beta - \omega t), \quad (4)$$

with \mathbf{k} being the wave vector of the surface wave, $k = 2\pi/\lambda$ the wave vector magnitude, β the wave angle of incidence with respect to the jaw's long axis, and ω the wave angular frequency. To solve the system (1)–(3) we insert a harmonic ansatz for the heave and pitch response,

$$u = a_u \cos(\epsilon_u - \omega t), \quad \theta = a_\theta \cos(\epsilon_\theta - \omega t) \quad (5)$$

where the amplitude and phase of the response are to be determined. The jaw movement is given by

$$h_{\text{jaw}} = u + (x - \bar{x})\theta. \quad (6)$$

Substituting (4) and (6) into (3) we see that every time derivative introduces a factor ω into (3). Since the frequency content of the incoming wave is peaked between 200 and 1000 Hz, the force equation will be strongly dominated by the inertia term and we obtain $f \approx m_H \ddot{h}$. Friction and buoyancy terms are small because the jaw-motion amplitude is much smaller than the jaw dimensions.

To make our estimate of the jaw-motion response explicit, we now assume a cylindrical shape for the two sides

of the lower jaw. We then have a constant cross-sectional mass density $m_H = m$, $M = mL$, and $J = mL^3/12$, with L as the length of the jaw. We take the COM to lie at the origin ($\bar{x} = 0$) so that the integrals in (1) and (2) run from $-L/2$ to $L/2$. Furthermore, we define $\xi := (kL \cos\beta)/2$. Substituting everything into (1) and (2) we find

$$\frac{a_u}{a_{\text{sand}}} \cos(\epsilon_u - \omega t) = \frac{\sin\xi}{2\xi} \cos(\omega t), \quad (7)$$

$$\frac{a_\theta}{a_{\text{sand}}} \cos(\epsilon_\theta - \omega t) = \frac{3(\sin\xi - \xi \cos\xi)}{L\xi^2} \sin(\omega t).$$

Since the above equations must hold for all times t we get for the amplitudes and phases

$$\frac{a_u}{a_{\text{sand}}} = \frac{\sin\xi}{2\xi}, \quad \frac{a_\theta}{a_{\text{sand}}} = \frac{3(\sin\xi - \xi \cos\xi)}{L\xi^2}, \quad (8)$$

$$\epsilon_u = 0, \quad \epsilon_\theta = \pi/2.$$

We can now calculate the deflection of the cylinder tip where $x = -L/2$, use (6), and find for its amplitude

$$a_{\text{jaw}}/a_{\text{sand}} = (a_u^2 + a_\theta^2)^{1/2}/a_{\text{sand}} \approx 1/2 + \xi^2/6, \quad (9)$$

where we have expanded the amplitudes in powers of ξ . Typical parameter values are given in Table I. For these values, $|\xi|$ lies between 0 and 0.6, depending on the input angle β , so that an expansion in terms of ξ is permitted. From (9) we see that the lower-jaw amplitude would be about half that of the incoming surface wave. Since the auditory system is sensitive to jaw movement down to \AA amplitudes, prey detection using the jaw-quadrates-stapes pathway would be not only possible, but even efficient.

Once auditory information has been processed by the cochlear hair cells in left and right ear, with a given interaural time difference (ITD) determined by prey direction β , their phase-locked spikes are sent to the *Torus semicircularis*, where signals from left and right ear come together. A reasonable mechanism for sound localization is constructing a neuronal map [12] of auditory space based on ITD cues, as suggested by Jeffress [13]. This scheme is attractive because of its simplicity and experimental evidence provided by multiple groups of terrestrial vertebrates, including other diapsid reptiles and their evolutionary descendants [14]. In the Jeffress scheme the auditory input from both cochleas is relayed to a set of map neurons along axonal delay lines (Fig. 4).

The map neurons are topographically organized. If a stimulus ITD at the ears is exactly compensated by the axonal delay difference, the map neuron receives *simultaneous* input from both ears and will fire at a maximal rate. If its delay does not match the input ITD, the neuron does not receive coincident input and its firing rate is low(er). Consequently every map neuron is tuned to a *specific* ITD and thus to a specific input direction. The bilateral pathways between ears and map neurons have fixed axonal delays that can effectively be ignored [12].

TABLE I. Parameters used in both analytical calculations and computer simulations. There is robustness against variation. Time constants are typical for the auditory system [11].

| Parameter | Value |
|------------------------------|--|
| Physical parameters | |
| Rayleigh wave speed | $v_{\text{Ray}} = 45 \text{ m/s}$ |
| Wavelength | $\lambda = 15 \text{ cm}$ |
| Input angular frequency | $\omega = 2\pi \times 300 \text{ Hz}$ |
| Jaw length | $L = 3 \text{ cm}$ |
| Inter ear distance | $d = 3 \text{ cm}$ |
| Maximal input ITD | $\text{ITD}_{\text{max}} = 667 \text{ }\mu\text{s}$ |
| Simulation parameters | |
| Number of cochlea neurons | $N_{\text{cochlea}} = 75$ |
| Cochlear firing rate | $A = 250 \text{ Hz}$ |
| Number of map neurons | $N_{\text{map}} = 100$ |
| ITD range of map neurons | $\text{ITD}_{\text{map}} \in [-1.33, 1.33] \text{ ms}$ |
| Synaptic strength | $J = 0.016$ |
| Post-synaptic current width | $\tau_{\text{EPSC}} = 250 \text{ }\mu\text{s}$ |
| Map neuron time constant | $\tau_{\text{relax}} = 500 \text{ }\mu\text{s}$ |
| Map neuron refraction time | $\tau_{\text{refr}} = 1 \text{ ms}$ |
| Map neuron capacitance | $C = 1$ |
| Map neuron resting potential | $V_0 = 0$ |
| Map neuron threshold | $V_{\text{thresh}} = 1$ |

We have numerically modeled this neuronal network using a right and left cochlear neuron population and a population of map neurons connected to them by axonal delay lines. Because of cochlear decomposition, cochlear neurons fire periodically. The only difference between left (L) and right (R) population is the phase shift due to the “stereo” delay. As soon as the cochlear neurons fire, all details about the mechanical input are lost, except for timing and accuracy of the spikes. We thus describe the firing of L and R cochlear neurons by an inhomogeneous Poisson process with rate function ($-\pi \leq \phi_{R/L} < \pi$)

$$\lambda_{R/L}(t) = \frac{A \exp[-(\phi_{R/L})^2/2\sigma^2]}{\sqrt{2\pi}\sigma \text{erf}(\pi/\sqrt{2}\sigma)}. \quad (10)$$

Here A determines the firing rate of a cochlear neuron. The input phase at right and left ear is given by

$$\phi_R = (\omega t)_{\text{mod } 2\pi}, \quad \phi_L = \left[\omega \left(t + \frac{d \sin\beta}{v_{\text{Ray}}} \right) \right]_{\text{mod } 2\pi}. \quad (11)$$

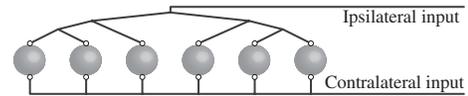


FIG. 4. (Pseudo-)Jeffress model of auditory localization. Neurons in the *Torus semicircularis* receive input with a fixed axonal delay from the ipsilateral ear and an axonal-distance-dependent delay from the contralateral ear. The ITD is determined by input direction. Through variable delays, different neurons are tuned to different directions [13].

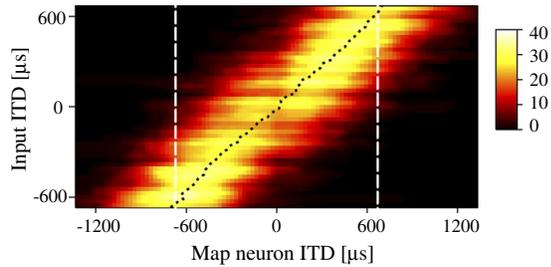


FIG. 5 (color online). Response of map neurons to input signals with $VS = 0.9$. Map neuron ITD is shown horizontally and input ITD vertically. Right-hand column shows the coding for number of spikes fired by the map neurons in a 250 ms time bin. White dashed lines are physical bounds to the input signal ITD, the black dotted line indicates the ITD estimate obtained by calculating the weighted mean of the map neuron responses. Clearly, a realistic estimate of input ITD is possible.

Equations (10) and (11) mean that, depending on the input phase, the firing probability peaks during every cycle of the incoming wave. The phase difference $\phi_R - \phi_L$, determining the timing difference between R and L ear, consists of the term $(d \sin \beta) / v_{\text{Ray}}$ (d = interaural distance, v_{Ray} = Rayleigh wave velocity), which gives the ITD between R and L depending on the angle of incidence β . The parameter σ in (10) determines the temporal accuracy (phase locking) of the cochlear neurons and can still be varied. A large value of σ causes broad peaks and implies imprecise firing whereas for a small value of σ the cochlear neurons respond very precisely to a periodic input signal. Phase locking quality is measured by the *vector strength* (VS) defined by $VS = |f^1|/f^0$, where f^n is the n th Fourier component of $\lambda_{R/L}$ in (10). The vector strength measures the sharpness of the response peak, but does not depend on the phase where the peak occurs. It therefore has an equal value for the R and L populations. By tuning σ it is possible to give VS any value between 0 and 1.

The map neurons (Fig. 4) are modeled as leaky integrate-and-fire neurons [15] whose membrane voltage V is governed by $\dot{V}(t) = -[V(t) - V_0]/\tau_{\text{relax}} + I_{\text{in}}(t)/C$. Here V_0 is the resting potential, τ_{relax} the membrane relaxation time, C the membrane capacitance, and I_{in} the synaptic input current. If V reaches a threshold value V_{thres} the neuron fires and the potential is reset to V_0 . After a refractory period τ_{refr} the evolution of the potential resumes. The input current I_{in} to the map neurons consists of the sum of excitatory postsynaptic currents (EPSCs) from the cochlear neurons; that is, from (10) and (11). Every input spike arriving at a synapse between a cochlear neuron and a map neuron leads to an EPSC of the form $I_{\text{EPSC}} = Jt/\tau_{\text{EPSC}}^2 \exp(-t/\tau_{\text{EPSC}})$ where τ_{EPSC} is the time constant of the EPSC, $t > 0$, and J characterizes the synaptic transmission strength.

Simulation results are shown in Fig. 5. For a realistic value of $VS = 0.9$ [16] the number of output spikes of the map neurons in a time interval of 250 ms is shown. The

firing rate is maximal if a map neuron's ITD corresponds to the ITD of the input signal but tuning is quite broad. Many map neurons encoding an ITD roughly matching that of the input respond with a fairly high firing rate. To estimate the input ITD from the firing rate of the map neurons the rate-weighted mean of the map neuron ITDS is calculated [17]. Such a neuronal mean, or *population code*, is effectively an actuator. This estimate comes close to the actual value of the input ITD. For $VS = 0.9$ the rms error (standard deviation) of the ITD estimate is $38 \mu\text{s}$, i.e., 3° in front of the animal.

In summary, we have modeled how snakes can use substrate vibrations to detect and localize prey. Our model of the two independent sides of a snake's lower jaw resting on the sand surface suggests that prey-generated waves can be perceived and localized reliably through a Jeffress setup. Furthermore, an estimate of the input signal ITD with microsecond accuracy is possible and high input vector strength is not needed. That is, we have provided strong support for the old hypothesis that snakes can hear, and thus use, substrate-based vibrations.

We thank the Zoologische Staatssammlung München, DFG, and BCCN—Munich for help and funding.

-
- [1] P.H. Hartline and H.W. Campbell, *Science* **163**, 1221 (1969).
 - [2] E. G. Wever, *The Reptile Ear: Its Structure and Function* (Princeton University Press, Princeton, NJ, 1978).
 - [3] B. A. Young, *Q. Rev. Biol.* **78**, 303 (2003).
 - [4] B. A. Young and M. Morain, *J. Exp. Biol.* **205**, 661 (2002).
 - [5] P. H. Hartline, *J. Exp. Biol.* **54**, 349 (1971).
 - [6] J. J. Thomsen, *J. Sound Vib.* **260**, 117 (2003).
 - [7] P. H. Brownell, *Science* **197**, 479 (1977); B. Aicher and J. Tautz, *J. Comp. Physiol. A* **166**, 345 (1990).
 - [8] K. F. Graff, *Wave Motion in Elastic Solids* (Oxford University Press, Oxford, 1975); D. Lohse *et al.*, *Phys. Rev. Lett.* **93**, 198003 (2004).
 - [9] I. S. Aranson and L. S. Tsimring, *Rev. Mod. Phys.* **78**, 641 (2006), Sec. III.B.
 - [10] J. N. Newman, *Marine Hydrodynamics* (MIT Press, Cambridge, MA, 1986).
 - [11] A. D. Reyes *et al.*, *J. Neurosci.* **16**, 993 (1996).
 - [12] E. I. Knudsen *et al.*, *Annu. Rev. Neurosci.* **10**, 41 (1987); R. Kempter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4166 (2001).
 - [13] L. A. Jeffress, *J. Comp. Physiol. Psychol.* **41**, 35 (1948); B. Grothe *et al.*, in *Evolution of the Vertebrate Auditory System*, edited by G. A. Manley, A. N. Popper, and R. R. Fay (Springer, New York, 2004), p. 289, Fig. 10.5.
 - [14] P. X. Joris *et al.*, *Neuron* **21**, 1235 (1998); K. M. MacLeod *et al.*, *J. Comp. Neurol.* **495**, 185 (2006).
 - [15] W. Gerstner and W. Kistler, *Spiking Neuron Models* (Cambridge University Press, Cambridge, 2002).
 - [16] G. A. Manley *et al.*, *J. Comp. Physiol. A* **167**, 129 (1990); C. Köppl, *J. Neurosci.* **17**, 3312 (1997).
 - [17] A. Pouget *et al.*, *Annu. Rev. Neurosci.* **26**, 381 (2003).