Sound Production and Hearing in Diverse Animals

Neville H. Fletcher
CSIRO Institute of Physical Sciences
Limestone Avenue
Canberra ACT

ABSTRACT: The vocal signals by which communication is achieved between members of the same species show remarkable similarity for animals as diverse as cicadas, frogs and humans. Likewise the auditory systems of these animals, although superficially very different, are all variations upon a single basic design. Some of these matters are discussed in detail and an explicit scheme for the analysis of auditory systems is given.

INTRODUCTION

Although modern acoustics is very much concerned with the behaviour of the human auditory system and with human vocal organs, it is not often that we stand back and take a much more generalised look at hearing and sound production in a large class of different animals. Certainly much of the classic work on human auditory systems relied heavily on studies of cats and guinea pigs, but in those cases similarity is quite close. What I would like to do here is to explore, much more generally, hearing and sound production in animals as diverse as humans, frogs, cicadas and crickets. I will take a physicist's approach and ignore biological niceties in order to present a coherent picture.

First let us note a few unifying principles which seem to apply to nearly all animals. The first is that sound production and hearing have developed essentially as a means of communication between members of the same species. Of course hearing also serves to give warning of the approach of predators or other dangers and vocal sounds can be used to frighten off intruders of other species, but that is not the primary evolutionary purpose of these physiological abilities.

The consequences of this communication function among members of the same species are several. In particular it leads to vocal and auditory capacities which are closely matched in frequency bandwidth and in time resolution. It also requires that each species develop a coding in the emitted sound which makes recognition of the species simple, and a further coding which conveys efficiently the desired information. There is remarkably little variation in the way in which this is done over the whole range of vocal animals.

Further, since a primary purpose of communication is social contact, all auditory systems possess some form of directional discrimination as well as an obvious intensity discrimination. At first sight the principles used appear to vary widely from one animal type to another, but we shall see that, rather surprisingly, most systems are actually variants of a particular generalised acoustical design with some features exaggerated and others suppressed.

SOUND PRODUCTION

For our present purposes, animals — and we are thinking here of land-dwelling animals — can be divided into two classes: those with lungs and those without lungs. All the larger animals, and in particular all the vertebrates, fall into the first class, while small animals such as insects may belong to the second class, which obtains its bodily oxygen needs by slightly forced diffusion along a series of tubes let into the side of the body.

An animal with lungs, though it may make sounds in a variety of ways, always produces its primary vocal sounds pneumatically by using air released from the lungs under pressure. Animals without lungs do not have this option and must produce sound by muscularly excited vibration. Let us discuss this second class first.

There are two common ways in which sound can be excited efficiently by muscular effort. Each involves a resonant structure and an exciting mechanism. In the first strategy the exciting mechanism is a sharp pic which is drawn across the ridges of a file-like structure to produce a train of regularly spaced excitation pulses. If the pic, or more usually the file, is closely coupled mechanically to a plate or membrane with a resonance frequency equal to that produced by the motion of the pic across the file, then a vibration of large amplitude will be generated and, if the vibrating structure is not too small compared with the wavelength of sound in air, efficient sound production will occur.

This mechanism, which is used by crickets and similar insects, already meets the communication needs of the animal. The song is based on a well-defined carrier frequency, that of the plate resonance, and consists of "syllables", one for each scrape of the pic across the file, as generated by a wing or leg motion. The information content of the song can be varied by changing the number of syllables in a "word" or by changing the repetition rate of syllables.

The other common mechanism, used by cicadas and similar insects, involves a ridged plate or tymbal closing the top of an air-filled cavity to produce a simple resonant structure. If the plate is distorted by muscular effort then it buckles progressively in a series of pulses related to the ridge spacing. If the pulse rate is equal to the cavity resonance frequency then sound is generated efficiently. The same coding possibilities exist. A typical example is shown in Figure 1.

For obvious reasons of efficiency both these mechanisms suggest the use of progressively higher frequencies for smaller insects, and this is what is observed in practice. For insects a few centimetres long, the song carrier frequency is typically in the range 3 to 5 kHz, and the syllable repetition rate perhaps 200 Hz. A remarkable amount of energy is often put into the production of the song — a continuously singing cicada may typically radiate 1 mW of acoustic power, which is comparable with the power of the human voice.

In insects the song serves primarily as a mating call and only the males have vocal organs. An interesting extreme case is the green Australian hedge cicada *Cystosoma Saundersii*. The male of the species has an enormous air-filled abdomen, about 4 cm long, which serves entirely as a resonator for the song, which has the unusually low frequency of 800 Hz with a 40 Hz syllable rate.

When we examine the production of sound by pneumatic power we find that just one mechanism is used. If the air passage is obstructed by a flap of cartilage, or by a pair of such flaps closing together, arranged in such a way that air pressure in the lungs will blow them open, then this simple system constitutes a pneumatic oscillator. Its operation is exactly similar to that of the lips in blowing a "raspberry", and the frequency of oscillation is just slightly above the natural resonance frequency of the vibrating flaps. The existence of a cavity on the air supply side of the system is crucial to the operation of the oscillator, as is the direction of opening of the flaps — a system with a flap blown closed like the reed of a clarinet operates in an entirely different and, for the present purpose, unsuitable way.

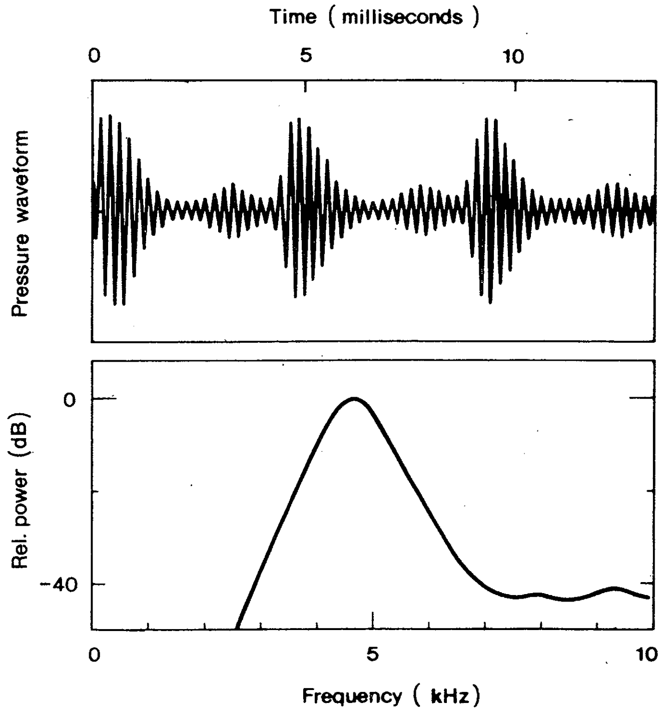


Figure 1: A typical insect song consists of a repeated series of "syllables", each associated with a wing or leg motion or with the contraction of a tymbal muscle, as shown here for the cicada. Syllables may follow each other continuously, as for the cicada, or may be grouped into "words" of three or more syllables. In more complex animals, such as frogs, the length of individual syllables in a word may vary. The acoustic spectrum is sharply peaked around the carrier frequency characteristics of the species, about 4.5 kHz for the cicada shown here.

This pneumatic oscillator is called a larynx, and systems of this type are found in humans, dogs, frogs and even (though in a modified and dual form) in birds. Although the larynx motion may be not too far from sinusoidal, the air flow through it has a strongly pulsed character and produces an excitation rich in harmonics.

It is clear that, once again, we have a vocal sound based upon a more or less regular carrier frequency, the basic oscillation frequency of the larynx. In humans this frequency itself carries valuable information — it is around 150 Hz for adult men, around 300 Hz for women, and 400 Hz or more for children.

Human speech obviously has a rich acoustic structure. Not only is the larynx sound broken up into syllables and words by changing air flow, but the distribution of energy in the upper parts of the rich pulse spectrum is varied by manipulating the frequencies of the first three resonances of the pipe-like vocal tract leading to the open mouth, by changing lip, tongue and mouth shape. The "formant" resonances are nominally around 500, 1500 and 2500 Hz, as we would expect for the impedance maxima of a cylindrical tube about 15 cm long, but can be shifted quite significantly to produce different vowels as shown in Figure 2. Human speech also involves assorted hisses and

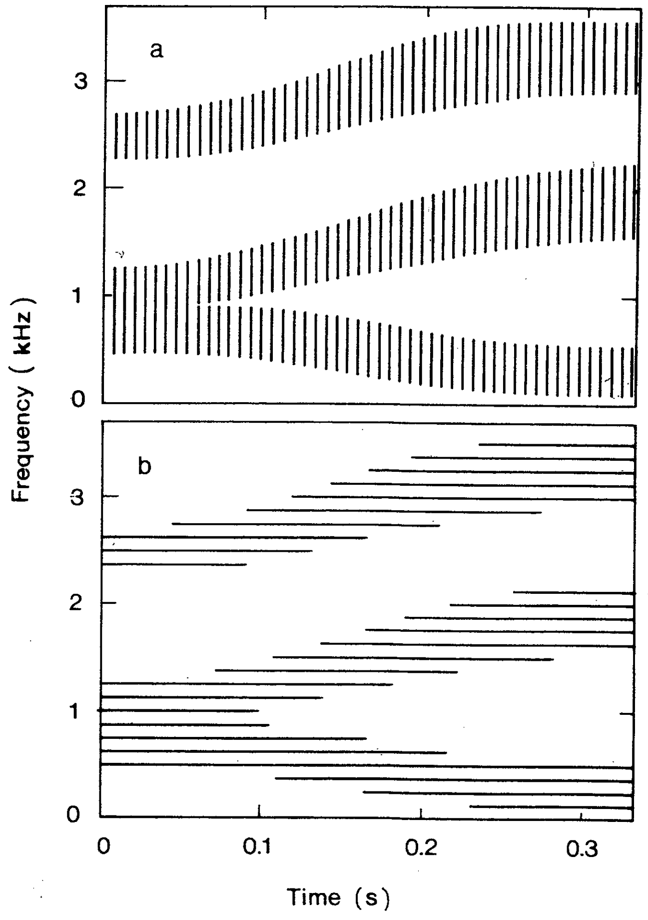


Figure 2(a): A Sonagraph-type display of the formant structure of a human speech diphthong "O-EE". The Sonagraph gives a frequency-time display with relative energy in different spectral ranges coded as blackness. Typically the Sonagraph filter has a bandwidth of 300 Hz so that individual frequency components of male speech cannot be separated, but the individual pulses from the larynx are resolved in time. The shift in frequency of the three major speech formants can be clearly seen.

Figure 2(b): If the bandwidth of the analysing filter is made much narrower, then time resolution is traded off for frequency resolution. Pulses from the larynx are no longer resolved in time but the individual harmonics of the pulse frequency are clearly visible. These vary considerably in intensity but to only a minor extent in frequency in normal speech.

clicks but the important part of the speech spectrum runs from about 300 Hz to 3 kHz. There is not a great deal of radiated energy at the larynx fundamental and in any case it is possible to deduce what its frequency is from the regular frequency spacing between successive harmonics.

Once again there is a correlation between larynx frequency and size in different animals but, since vertebrates change size significantly during their lives, this frequency is not of dominant importance for species recognition. Instead, reliance is placed upon other features of the rich vocal sound. Some more primitive animals, however, have quite sophisticated "built-in" coding. An example is the frog, some species of which have two sets of vocal flaps, in series, one with a frequency around 100 Hz and one with a much higher frequency, say 2 kHz, giving a song sound quite closely related to that of an insect.

In humans, of course, the larynx frequency can be varied by a factor of 4 or more in the artificial activity of singing, while birds do this as part of their normal vocalisation. Other pneumatically generated noises such as whistles, which are generated by the interaction between an unstable air jet and a resonant cavity, can be used as part of a song or speech but these are generally not the basic speech sounds of a species.

AUDITORY SYSTEMS

The basic requirement of an auditory system is that it be sensitive to sound pressure over the frequency range characteristic of the song or speech of the species and that its internal mechanism convert the air pressure variations associated with the sound into mechanical displacement of some internal structure to which nerve-cell transducers are attached.

Some animals, such as caterpillars, have sensory hairs which perform a similar function in relation to the acoustic velocity signals produced by the beating wings of predators, but essentially all genuine auditory systems seem to rely upon taut diaphragms as the primary transduction mechanism from acoustic pressure to mechanical vibration. The diaphragm or tympanum may be assisted, as we see later, by various resonators, couplers or phase shifting networks depending upon particular features that evolutionary processes have emphasised, while the complexity of the mechanical-to-neural transducer varies greatly with the sophistication of the vocal code of the animal concerned.

The upper part of Figure 3 shows a generalised auditory system from which we can consider all real animal auditory systems to be derived by suppressing one or more features. Interestingly all the features shown are present in the auditory systems of higher mammals, including humans, but the Eustachian tubes which connect the cavity behind the tympana to the rest of the system are so narrow relative to the size of the tympana that effectively no sound energy travels along them and the system behaves like two isolated simple ears. In this and all other systems we assume that a neural transducer is connected to each tympanum to convert its motion to nerve impulses, but we forego any consideration of detail.

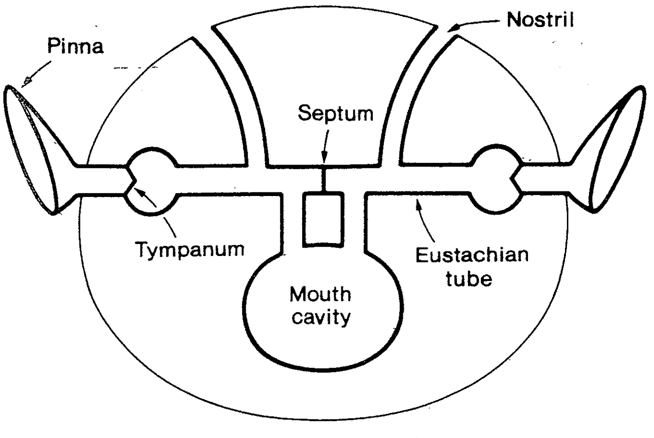


Figure 3: A generalised model auditory system from which most other systems can be derived. In mammals the Eustachian tube is so narrow that the system functions as two independent ears. In simpler animals there is generally an acoustic coupling between the two ears but some other features of the generalised system may be missing.

The isolated ear of the mammalian auditory system has a simple acoustic function, at least in outline, though its performance is rich in detail at higher frequencies. The pinna provides an acoustic pressure gain of 10 to 20 dB at frequencies above a few kilohertz, depending on shape and size, but the gain falls to not much above 0 dB below a few hundred hertz because of the transmission behaviour of finite flared horns. The tympanum itself is typically a tensioned conical membrane with a resonance frequency around 2 kHz, near the middle of the hearing range of interest. The Q value is typically less than unity, say about 0.5, so that its frequency response is broad. Taking account of the performance of the horn-like pinna we therefore have a system with a fairly flat frequency response from around 1 kHz to 5 kHz, with a decreasing response below 1 kHz because of the combined effects of horn cut-off and tympanum resonance, and a decreasing response above about 5 kHz because this is well above the tympanum resonance. Actually the sensitivity of the human ear is much less than would be predicted by this model below 100 Hz and above 10 Hz but this is undoubtedly due to evolutionary limitations on the neural transduction mechanism rather than to simple acoustic response.

The directional discrimination of a single mammalian ear is influenced both by the geometry of the pinna, which tends to emphasise high frequency sounds incident from along the direction of its axis, and by diffraction around the head, which produces directional maxima and minima in a more complicated way, though again favouring incidence from the ear direction. On top of this there are other complications caused by the convoluted shape of a typical mammalian pinna and, of course, the subtle psychophysical correlations between the inputs of two independent ears that are made possible by the sophistication of the mammalian brain.

COUPLED EARS

In all the other animals I will discuss here, the acoustic pathways between the two ears are sufficiently open that we must consider them in combination. It is interesting that this is the situation in the lower animals whose response to auditory stimuli is much simpler than in mammals. By evolutionary chance it proved more appropriate to rely upon increased acoustic sophistication at the auditory periphery rather than to undertake more complex neural processing.

The simplest auditory system is that belonging to the frog, as shown in Figure 4(a). The tympana are simple extensions of the outside skin of the head, and short wide Eustachean tubes lead directly from the ears to the mouth cavity. By chance the geometry of the auditory system of insects like the cicada, as shown in Figure 4(b), is very similar, though in this case the two tympana and the cavity are located in the abdomen and the cavity has no other function.

Acoustic analysis of this sort of system is quite straightforward, both in relation to frequency response and to directional discrimination. While diffraction around the body of the animal will certainly have some effect, this is relatively minor compared with the effect of phase differences between the sound pressure at the two tympana for various incidence directions.

Since the cavity is reasonably large, though small in dimensions relative to the sound wavelength involved, it behaves as a simple acoustic compliance, and the tympana have maximal response near their resonance frequency, which we would expect to be tuned to the carrier frequency of the song of the species, about 1.5 kHz and about 5 kHz respectively for these two cases.

It turns out from the analysis that two ears coupled by a cavity in this manner can have a cardioid response with a directional discrimination of nearly 20 dB if the Q value for the tympana is appropriately chosen. For a typical frog-like case the appropriate value is around 4, which implies that the bandwidth of the system will not be large. This is, of course, exactly what is required, given the simple amplitude-modulated nature of the frog song. An exactly similar conclusion is reached in the case of the cicada.

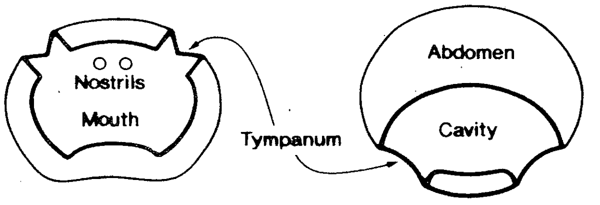
Since the primary purpose of the song of both male frogs and cicadas is to attract females while keeping males of the same species at a distance, this directional discrimination and frequency specificity has obvious behavioural importance. All the female has to do is to respond to the carrier frequency, check the modulation identification, and then hop in the general direction of the ear giving the stronger signal.

Birds are clearly more sophisticated animals than frogs, and most species have a complicated song which covers a reasonably wide frequency range, generally starting above about 500 Hz. It is interesting, therefore, to note the rather simple acoustics of the auditory system, a design shared with some reptiles, thus indicating their common ancestry. This is shown schematically in Figure 4(c).

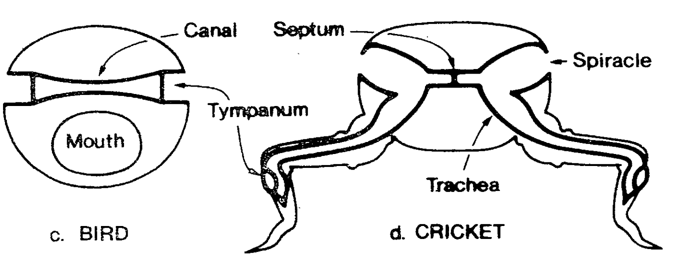
At the frequencies involved in birdsong, the length of the auditory canal is not small compared with the wavelength of sound and it is therefore necessary to be rather more sophisticated about the acoustic analysis. Not only must wave propagation along the canal be considered, but also attenuation of those waves through wall effects.

Once again it turns out that the response for the ear facing the sound source is greatest near the resonance frequency, but it also tends to be large near the frequency at which the canal length is one quarter of a wavelength, so that if these two frequencies agree the effect is maximal. It also happens that, at this special frequency, the response of the ear facing away from the sound can actually be brought to zero by proper choice of the loss in the canal and the Q value of the tympana. Around this frequency the response is cardioid while for higher or lower frequencies the directional pattern is less asymmetric. Of course we would not expect a real system to behave in exactly this way, but the qualitative behaviour should be very similar to that calculated.

Finally we see in Figure 4(d) a representation of the auditory system of the cricket. As in mammals it is associated with the respiratory system but, since the cricket is an insect, this system has a very different structure. The auditory system is located in the first pair of legs of the animal. Each leg contains a large tracheal tube connecting to a more or less open spiracle in the prothorax, and this trachea has a thin-walled tympanum (or rather a pair of tympana) in the leg wall just below the knee. The two tracheae are connected by a short tube which generally has a central septum.



a. FROG b. CICADA



c. BIRD d. CRICKET

Figure 4: Specialised forms of the general model auditory system. (a) in the frog the tympana connect directly to the mouth cavity through short Eustachian tubes; (b) the cicada has a similar system but the cavity is in the abdomen; (c) the bird has a simple constricted canal connecting the two tympana; while (d) the cricket has a tubular tracheal system with tympana on the legs and open spiracles, the geometry of which varies from species to species.

Different species of cricket have different aspects of this geometry exaggerated. In some the spiracles are large and wide open and the tracheae have the form of flaring internal horns with little interconnection. In other species the tracheae are nearly cylindrical, the spiracles are small and partly covered, and the interconnecting tube is short and wide with a thin septum. Analysis of such systems with their multiple ports for sound entry and multitude of unknown parameters is a complex business. In particular the narrow tracheae have considerable wall loss and, since the sound wavelength at 5 kHz is comparable with the size of the animal, the exact placing of the legs can be important. Again however it seems clear that this system can provide substantial directionality of response.

As in man, it turns out that the neural response of many of these animals is much more restricted in frequency than we would expect from simple acoustic analysis. Such neural frequency selectivity is by no means unexpected but its mechanism is as yet obscure to those neurophysiologists studying the problem.

CONCLUSION

We have seen that systems for sound production and hearing in the animal world appear to have a common origin and a common purpose, though the ways in which they have developed show a great diversity. Understanding the acoustics of these systems is, of course, only the beginning of the study, and by far the greatest interest lies in the neurophysiological and psychophysical aspects of the problem. Nevertheless an examination of simple acoustics does give us an insight into the way in which these systems function, and acoustics provides an essential framework around which more sophisticated studies can be built.

BIBLIOGRAPHY

For those who would like to read further about the subject, the references below provide more detail on some of the things discussed in the paper and also copious references to the current literature.

H.C. Bennet-Clark, "Acoustics of insect song", Nature, Lond. *234*, 255-259 (1971).

N.H. Fletcher, "Air flow and sound generation in musical wind instruments", Ann. Revs. Fluid Mech. *11*, 123-146 (1979).

N.H. Fletcher & K.G. Hill, "Acoustics of sound production and hearing in the bladder cicada *Cystosoma Saundersii* (Westwood)", J.Exp. Biol. *72*, 43-55 (1978).

N.H. Fletcher & S. Thwaites, "Physical models for the analysis of acoustical systems in biology", Quart.Rev.Biophys. *12*, 26-66 and 463 (1979).

N.H. Fletcher & S. Thwaites, "Acoustical analysis of the auditory system of the cricket *Teleogryllus commodus* (Walker)", J.Acoust. Soc.Am. *66*, 350-357 (1979).

W.L. Gulick, "Hearing: Physiology and Psychophysics", Oxford University Press, New York (1971).

B. Lewis (editor), "Bioacoustics: A Comparative Approach", Academic Press, London (1983).

A. Michelsen & H. Nocke, "Biophysical aspects of sound communication in insects", Adv. Insect Physiol. *10*, 247-296 (1974).

G. Sales & D. Pye, "Ultrasonic Communication by Animals", Chapman & Hall, London (1974).

E.A.G. Shaw, "The external ear" in "Handbook of Sensory Physiology" Vol. V(1) ed. W.D. Kneidel & W.D. Neff, Springer-Verlag, New York 1974.

J. Sundberg, "Acoustics of the singing voice", Scientific American 82-91 (March 1977).

R. Teranishi & E.A.G. Shaw, "External-ear acoustic models with simple geometry", J.Acoust.Soc.Am. *44*, 257-263 (1967).

APPENDIX

For those interested in more detail, the analysis of frequency response and directionality for two simply coupled ears can be set out in a few lines. We choose a system like that of the bird and idealise it to a model with two identical tympana coupled by a uniform tube. The acoustic behaviour of each tympanum is expressed by its acoustic impedance Z_T — the ratio of acoustic pressure p to acoustic volume flow U at a given frequency ω. For a tympanum of mass M, area A, resonant frequency ω_0 and damping R

$$Z_T = R + j(M/A^2)(\omega - \omega_0^2/\omega). \qquad (1)$$

Description of the behaviour of the tube requires consideration of quantities p_1, U_1 and p_2, U_2 at its two ends and, if we choose the positive directions for both U_1 and U_2 to be into the tube, we can write

$$p_1 = Z_{11}U_1 + Z_{12}U_2 \qquad (2)$$

$$p_2 = Z_{21}U_1 + Z_{22}U_2. \qquad (3)$$

For a tube of cross section S and length L it can be shown that

$$Z_{11} = Z_{22} = (\varrho c/S)\coth[(\alpha + j\omega/c)L] \qquad (4)$$

$$Z_{12} = Z_{21} = (\varrho c/S)\,\mathrm{cosech}[(\alpha + j\omega/c)L] \qquad (5)$$

where ϱ is the air density and c the velocity of sound in air, and α is the attenuation coefficient, due to wall effects, for sound propagating along the tube.

Now we put the tympana across the ends of the tube, noting that this means that the acoustic flow through each tympanum is equal to that into the associated tube end. If p_L, U_L and p_R, U_R are acoustic quantities for the left and right ears respectively, then

$$p_L = (Z_{11} + Z_T)U_L + Z_{12}U_R \qquad (6)$$

$$p_R = Z_{21}U_L + (Z_{22} + Z_T)U_R. \qquad (7)$$

If sound comes from a direction at an angle θ to the left of straight ahead, then there is a phase delay at the right ear given by

$$p_R = p_L \exp[-j(\omega L/c)\sin\theta] \qquad (8)$$

and, using the symmetry relations in (4) and (5) in (6) and (7), we easily find the response of the left ear to be

$$\frac{U_L}{p_L} = \frac{(Z_T + Z_1) - Z_{12}\exp[-j(\omega L/c)\sin\theta]}{(Z_T + Z_{11})^2 - Z^2_{12}}. \qquad (9)$$

The linear motion of the left tympanum is given simply by

$$x_L = -jU_L/\omega A. \qquad (10)$$

This response is greatest at the frequency for which the denominator of (9) is smallest, which is a compromise between the resonance frequency of the tympanum ($Z_T \rightarrow R$) and the half-wave resonance of the tube ($Z_{11} \rightarrow Z_{12} \rightarrow (\varrho c/S)\,\mathrm{cosech}\,\alpha L$). Directional response is greatest for sound from the left ($\theta = 90°$) and can be made to vanish at one particular frequency for sound from the right ($\theta = -90°$) by appropriate choice of the damping coefficients R and α. In general there will be a response minimum for the left ear for sound coming from somewhere to the right of the animal ($\theta < 0$) but this will not be an exact null. Figure 5 shows calculations for a realistic case. The system therefore shows very useful frequency discrimination and directionality.

Exactly the same analysis can be used for cavity-coupled ears, such as those of cicadas and frogs, at frequencies for which the distance between the ears is only a small fraction of a wavelength. In this case

$$Z_{11} \approx Z_{12} \approx -j\varrho c^2/V\omega$$

where V is the volume of the cavity. The frequency response and directionality are similar to those of the tube coupled ears provided the cavity volume is appropriately chosen.

(Received 3 May 1985)

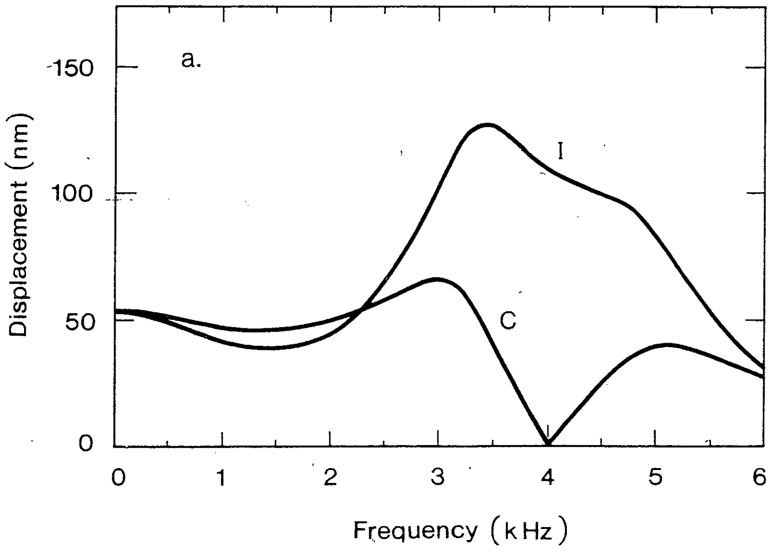


Figure 5(a): Calculated frequency response for a tube-coupled auditory system, with tube length 20 mm, tympanum resonance frequency 4 kHz and optimal damping, for ipsilateral sound I ($\theta = 90°$) and contralateral sound C ($\theta = -90°$) at 91 dB.

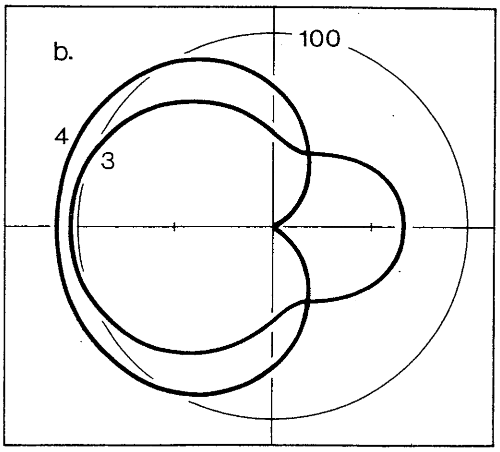


Figure 5(b): Calculated polar response of the left ear at 3 kHz and 4 kHz. The straight-ahead direction ($\theta = 0$) is upwards on the page. [From Fletcher & Thwaites 1979]