Singing and Song: A Biophysical and Evolutionary Perspective

Dianne Cameron

Laurentian University, Sudbury, ON

Abstract

The periodicity of natural signals from our environment since the dawn of life on earth has provided the foundation for development of rhythmic responses innate to living organisms, ranging from ancient bacteria to modern humans. The evolution of organs tuned to auditory signals – especially since the colonization of terrestrial environments by early vertebrates – has been matched by the coevolution of vocalization from the production of simple warnings and alarms to complex patterns communicating sophisticated messages. The need to interpret and produce auditory information is essential for a diverse spectrum of animals. At what point does complex vocalization become "song", and what is it that we regard as "singing"?

We recognize "songs" and "singing" in the spring chorus of amphibians, the humming and chirping of insects, the species-specific songs of birds, and the complex communication of humpback whales and other cetaceans. These terms vary in meaning when describing such vocalizations compared to humans. For human singers, the act of singing and the meaning of song have strong emotional and physical components that are profound and personal, with demonstrated effects on brain activity and development, health and well-being. Is this uniquely human?

This paper reviews the biophysics of hearing and the concept of singing from the perspective of evolution within the natural rhythmic soundscape of our world.

Keywords: biophysics of hearing, perceptual processes, evolution and paleontology, vertebrate transition to land, singing, song

The formative age of the earth was noisy and violent – but nonetheless filled with pulses and rhythmic fluctuations of light and sound. About 4.6 billion years ago (BYA), the earth was a molten surface with little atmosphere, no free oxygen (O_2) and no liquid water (H_2O). The earliest atmosphere was mostly hydrogen sulfide, methane and carbon dioxide (CO_2) from volcanic activity. It took about half a billion years before the surface was cool enough and solid enough for liquid water to form on it, and the dense greenhouse gases (up to 200 times current carbon dioxide levels) kept the temperature high.

The first life in the nascent oceans arose during the Archean period (4 – 2.5 BYA) as single cells using sulfur and other elements as their energy sources. By 2.7 BYA blue-green cyanobacteria had evolved to use the abundant CO_2 and water as energy sources through photosynthesis, releasing oxygen as a waste product into the atmosphere, until it eventually reached the present day level of 21%. These simple organisms were immersed in the noisy environment of cataclysmic eruptions and meteor impacts, and subjected to daily ocean tides and seasonal changes in periodicity of sound from wave and wind action. The sense of hearing and ultimately the concept of singing arose through the evolution of life within the natural rhythmic soundscape of our world.

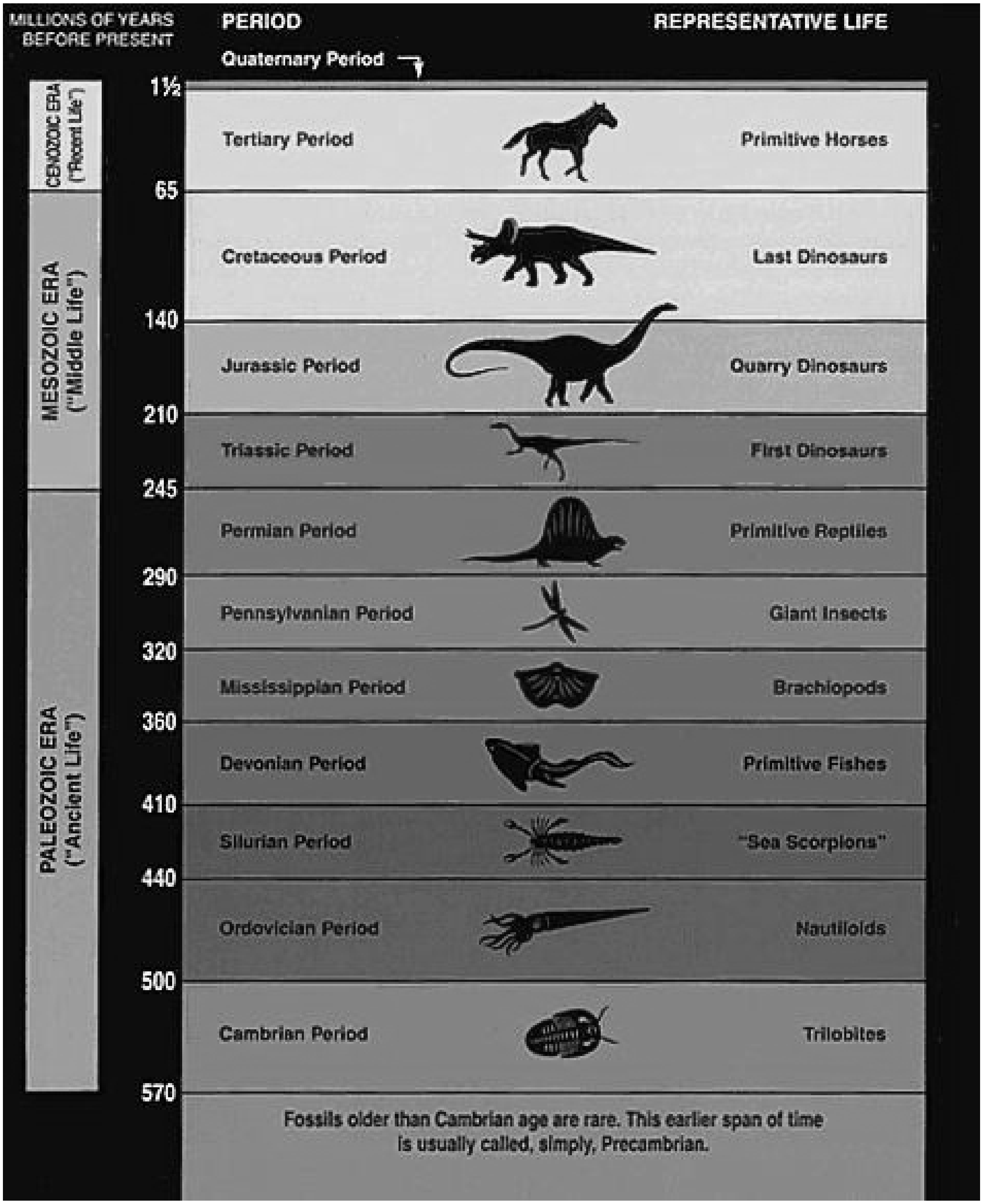


Figure 1. Highlights in geologic time. (Illustration from National Park Service)

With the most recent past at the top and the most ancient at the bottom, this figure presents the main divisions and subdivisions of geologic time in millions of years ago (MYA), illustrated by typical life forms of the period. The time before the Cambrian (not shown) is the Archean, from about 4 to 2.5 billion years ago (BYA), the age of Archaea and Bacteria (prokaryotes).

Major evolutionary events (Dorling Kindersley, 2012), from most recent to earliest, include:

- Cretaceous (140-60 MYA) to present: increased diversity of land animals, birds; singing and hearing
- Jurassic 200-140 MYA: 1st birds, frogs
- Triassic 250-200 MYA: 1st dinosaurs, turtles, mammals; advances in ear and vocal apparatus; 1st external ears
- Permian 300-250 MYA: rise of synapsids (mammal ancestors), evolution of ear bones; greatest mass extinction event
- Carboniferous 350-300 MYA: plant, insect, amphibian diversity; 1st reptiles; ear drums (Tympanum)
- Devonian 410-350 MYA: more land plants, insect progenitors; 1st amphibians, lung fish; 1st singing?
- Silurian 440-410 MYA: 1st jawed fish
- Ordovician 490-440 MYA: 1st land plants and arthropods; jawless fishes
- Cambrian 540-490 MYA: Cambrian “Explosion” of animal diversity; exoskeletons; 1st arthropods
- Proterozoic 2.5- 0.5 BYA: 1st eukaryotes (protozoa)
- 2.5 BYA: Great Oxygenation Event
- Archean 4-2.5 BYA: Archaea and Bacteria (prokaryotes)

Sound and Pressure in the Early Earth Environment

Though the early single-celled life forms had no ears, they nonetheless detected and responded to sound. “Sound” is the term for our perception of pressure fluctuations travelling as mechanical waves through the surrounding medium – air, water, or even solids. Sound waves are simply vibrations of the medium as minute pressure oscillations – kind of like a spring – which radiate away from the source of disturbance (or sound source) in all directions. The first cells detected pressure fluctuations in their watery environment, and evolved ways to interpret them.

Physical Descriptions of Sound

Sound transmits as a mechanical pressure wave that displaces particles of the medium through which it moves (air, water, solids) (Cameron, Skofronick, & Grant, 1999). Speed of transmission depends on density of the medium, slower in air and faster in solids. Hearing is thus a type of sensing of pressure. Oscillations of a typical sound wave are described in terms of repetition time (frequency and period), or distance between repetitions on the wave (wavelength): the number of cycles of repetition per second is the wave frequency, in Hz; period is the time it takes for a wave pattern to repeat, the inverse of frequency. “Pitch” is the perceptual term for frequency. Amplitude is a measure of how much the sound wave displaces particles of the medium as it oscillates. Velocity (v) of a sound wave (the speed with which it moves from its source through a medium), frequency (f), and wavelength (λ) are mathematically related: $v = f \lambda$

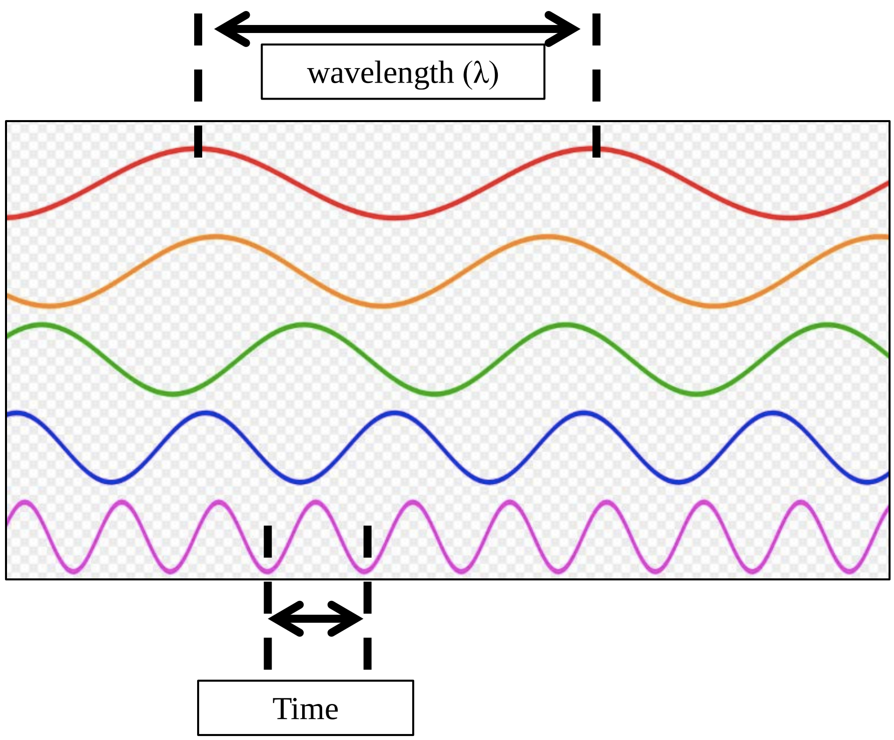


Figure 2. Describing a sound wave. (Illustration from Yoni Levinson, 2008)

Sound waves are typically described in terms of the repetition time (period) and distance along the wave (wavelength λ) between identical pattern repetitions of the travelling wave. Here waves of different frequencies, wavelengths and periods are illustrated; longest period and wavelength waves are at the top, and have the lowest frequencies (pitches).

Sound Intensity and Loudness

Intensity of a sound is the rate of energy flow (i.e. sound power) per unit area on which it falls. As we know from our own experience listening to music close to a sound source or far away from one, intensity decreases with distance from a sound source. Higher intensity sounds have larger amplitudes of displacement and larger oscillations in pressure – again, our personal experience demonstrates that a high intensity sound, such as a jack hammer, produces enough pressure to be painful to the relatively small surface area of the ear drum on which it falls. The perceptual term for intensity is loudness. Intensity is measured in watts/metre^2; loudness is measured in decibels or phons (Cameron, Skofronick, & Grant, 1999).

Speed of Sound and Impedance

Sound waves move through a medium at a speed dependent on properties (e.g. density) of the medium (gases, liquids and solids):

- In air: about 340 m/s
- In fresh water: about 4x as much, around 1480 m/s
- In sea water: around 1500 m/s
- In bone: about 4000 m/

Acoustic Impedance

If any mechanical wave encounters a boundary between media of different densities (e.g. when a water wave hits a wall) some of the wave is reflected and some is transmitted: this is termed impedance. Acoustic impedance occurs when sound transmitted through air or water meets the boundary of the hearing organ (e.g. ear drum) of an animal. Density and speed differ in the media on either side of the boundary, so much of a sound is reflected at the ear drum while little is transmitted across to internal auditory organs

Sound and Hearing

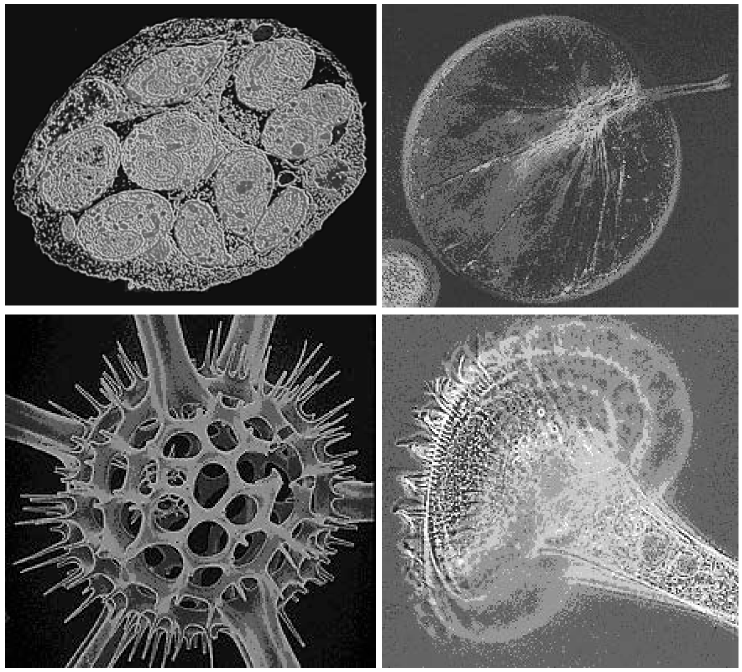
Hearing (or auditory perception) is the ability to perceive sound by detecting vibrations, the changes in pressure of the surrounding medium through time, through an organ such as the ear.

In humans and other vertebrates, hearing is performed primarily by the auditory system: sound waves detected by the ear are transduced into nerve impulses that are interpreted by the brain (primarily in the temporal lobe). Major adaptations were required in the transition of animals to land, since hearing in seawater had very different requirements from hearing in air. Due to acoustic impedance, most sound is reflected at the ear drum and little is transmitted across to internal auditory organs. Ways to compensate for this loss of sound evolved rapidly in land animals, along with ways to vocalize.

Origin of Sound Perception

Primitive cells (prokaryotes) such as bacteria and archaea (2-4 billion years ago) could sense pressure in their watery environment, but did little to interpret or use the sensation. A major advance came when single celled organisms found a new energy source: each other. Unlike the mostly unmoving bacteria and archea that used simple chemicals, or the cyanobacteria which photosynthesized, protozoa (about 750 MYA), were single celled motile organisms that advanced the sensing of pressure because they could move. Their motion created detectable minute pressure waves in the early ocean. Sensing these pressure fluctuations allowed protozoa:

- To detect prey
- To avoid predation
- To recognize and communicate with each other



Apicomplexan - moves by gliding

Flagellate - moves by whip like flagellum

Sarcodine - moves by pseudopods, includes amoebas

Ciliate - moves by many hair-like cilia

Figure 3. Eukaryotic Cells: Protozoa

Eukaryotic cells have a greater degree of internal organization, including a nucleus and a variety of specialized organelles. All multi-celled plants, fungi and animals are eukaryotes, but there are many more free-living single-celled and colonial forms of eukaryotes. Protists are by far the most diverse of all the kingdoms of living organisms. They live in all sorts of habitats and have many different specializations for feeding, locomotion and reproduction (Dorling Kindersley, 2012).

Origins of Singing and Song

Sensing of pressure in a water environment was the foundation for the sense of sound. Although invertebrates and fish can certainly produce sounds, the foundation for the first singing required the transition to land by arthropods, about 450 million years ago in the late Ordovician period.

Earth's first singers, the land arthropods, were ancestors of insects. Their transition to land resulted in major changes in both hearing and sound production. Ways to amplify sound arose to prevent loss (acoustic impedance) in the transition from air to the internal auditory structures: the ear drum, or tympanum. The first tympanic ear arose in insects, then in the first land vertebrates. Ways of funneling sound (flexible external ears) and filtering and amplification of certain pitches evolved. Ways of producing sound coevolved with ways of detecting sound. Among the many adaptations and evolutionary innovations, only a few examples are presented here.

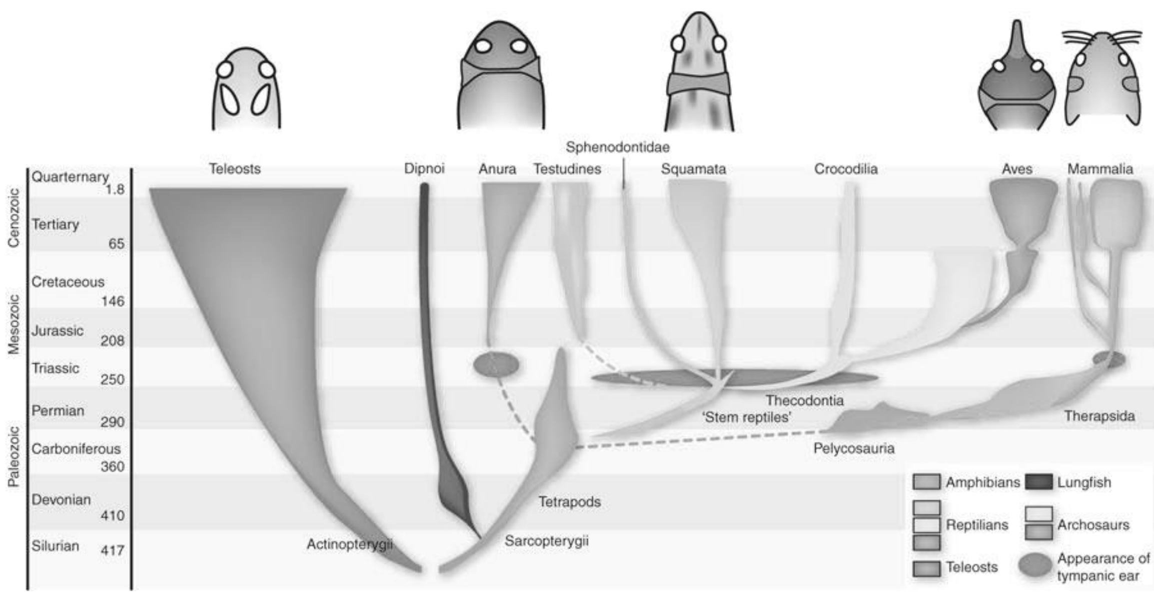


Figure 4. Evolution of the tympanic ear in vertebrates

The tympanic ear, a hearing organ with a drum-like membrane to amplify sound, first appeared in insects in the Devonian (about 400 MYA), then in various lineages of amphibians and reptiles through the Carboniferous (350 MYA), Permian, and Triassic (250 MYA).

Insect hearing: Insect Tympanic Ears

The tympanum helps amplify sound during transmission from air to the inner ear. Insects have oval eardrums, or tympana, a localized thinning of cuticle at the site of their hearing organ.

Crickets and katydids have tympana on the front legs at the base of the tibia; locusts have tympana covered by their wings on the sides of the first abdominal segment. Cicadas have exposed eardrums on the abdomen next to their tymbals. Male cicadas would deafen themselves producing such loud sounds right next to their ears, but have evolved an ingenious way to protect their hearing. As the male cicada starts to sing, he contracts a small muscle to fold the ear shut. In general, the tympana of singing insects are relatively insensitive to changes in pitch but are very sensitive to changes in the intensity of sounds being received. This corresponds to the basic structure of most insect songs, which rely more on variations in timing and pattern repetition than on changes in pitch (Hershberger, 2015).

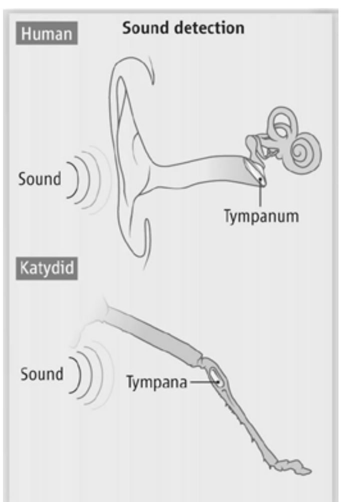


Figure 5. Tympanic ear in a human and a katydid.

Various singing insects have developed tympanic ears that are located and optimized differently in different species. In this example, the tympana are located on the front legs of katydids, but locust and cicada eardrums are on their abdomens. Vertebrate ears have other adaptations, such as the amplifying lever action of ossicles adjacent to the human tympanum, to further enhance sound detection.

Insect Song and Chorusing

Insects make sound in various ways: rubbing body parts against each other, tapping, or scraping. In crickets and katydids a sharp edged scraper on the upper surface of the lower wing is rubbed against a row of bumps (the “file”) on the underside of the upper wing. Grasshoppers use a similar method rubbing their hind legs. Male cicadas have special sound-producing organs or “tymbals” on the sides of the abdomen behind their wings. Contraction of muscles causes ribs in the tymbal to bend, producing sounds that resonate in the large tracheal air sac inside the abdomen. Cicadas produce the loudest of insect sounds.

Songs

Most insect songs are in the frequency range of 2,000 Hz to 15,000 Hz. Songs may be musical trills (continuous train of notes too fast to count, lasting several seconds) or chirps (short burst of notes lasting a fraction of a second), usually given in a series, each chirp being followed by a brief period of silence. Each species has its own distinct song, which is recognized by all individuals of the same species. Songs are told apart both by their dominant frequency and by details of their timing patterns (Hershberger, 2015).

Choruses

Some species form singing aggregations, in which males group together within appropriate habitat –e.g. cicadas. They are attracted to one another’s calls and form dense choruses. Calling often has contagious elements - the first male to begin singing in a group often elicits singing in other males. Singing insects may also synchronize their songs. The Common Meadow Katydid song is composed of a series of ticks followed by a buzz. Individuals in dense colonies often synchronize their songs so that they are all ticking together and then buzzing together. Snowy Tree Crickets often sing in unison, and it is not uncommon to find a number of males chirping together in almost perfect synchrony.

Milestones in the vertebrate transition (Carroll, 1988); (Allin, 1975); (Dorling Kindersley, 2012) to land include:

- Silurian 440-410 MYA: 1st jawed fish
- Devonian 410-350 MYA 1st land vertebrates
- Carboniferous 350-300 MYA: 1st reptiles; vertebrate ear drums (Tympanum)
- Permian 300-250 MYA: rise of synapsids (mammal ancestors), evolution of ear bones
- Triassic 250-200 MYA: 1st external ears

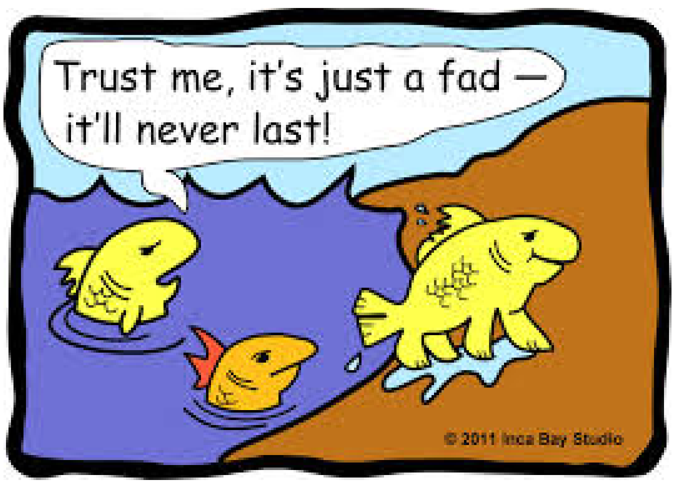


Figure 6. Vertebrate transition to land. (Illustration from Inca Bay Studio, 2011)

Sounds reaching the ear canal (auditory canal) are transmitted across the ear drum to the middle ear, where the levering action of the ear bones called ossicles amplifies the sound to compensate for acoustic impedance due to the change from air to tissue to fluid of the middle ear. Ear bones (ossicles) evolved from bones of the jaw. In early reptiles, the evolution of the ossicles took different paths. Reptiles and birds have only 1 ossicle, while mammals have 3 ossicles - the hammer (malleus), anvil (incus) and stirrup (stapes) (Allin, 1975); (Manley & Gleich, 1989).

More Adaptations for Hearing on Land

The first external ears, or pinnae, appeared in the Triassic about 250-200 MYA. External ears soon evolved many specialized shapes for funneling sound into the ear canal and detecting the direction of sound. They could also close off and protect the ear from loud sounds, and in many animals they serve as an efficient heat exchanger to prevent overheating. In addition, the combination of the length of flexible pinnae with the length of an animal's ear canal acts as a resonance chamber to selectively boost certain frequencies of sound, usually in the animal's vocal range, for enhanced communication within a species (Cameron, Skofronick, & Grant, 1999).

Figure 7. Localization and funneling of sound enhanced by flexible pinnae.

In the animals illustrated here, the flexible pinnae serve both as heat radiators for adaptation to hot environments, as well as effective funnels for capturing sound and discriminating which direction sounds are coming from.

Relationship Between Vocal Range and Frequency

The natural resonant frequency of the auditory canal depends on its length, just as the frequencies played on a whistle are related to the length within the tube between air input and air outlets. In animals with movable pinnae, the shape of the pinnae can funnel sound into the ear canal and also extend the effective length of the ear canal over which a resonant frequency can be created in a standing wave. This increases sensitivity at the resonant frequency – often in the middle of the animal's vocal range (Cameron, Skofronick, & Grant, 1999). The vocal range of a bat, with a very short ear canal, is very different from that of an elephant with a long ear canal, and reflects their specialized adaptations to particular environments and life styles. For example:

In humans:

- Outer ear (pinna) is not flexible, aids little in funneling sound
- Does not extend the effective length for resonance
- Ear canal length is about 2.5 cm, resonant frequency about 3400 Hz
- This falls in the frequency range of the human voice, roughly 250-5000 Hz

In elephants :

- Communicate over very long distances using low frequencies < 500 Hz and by infrasound (< 30 Hz) below the range of human hearing
- Ear canal of about 20 cm in length, large movable pinnae measuring about 180 cm at the widest
- Can extend the resonant frequency of the ear canal from 425 Hz down to 42.5 Hz in the very low frequency range used to communicate over distances up to 4 km.

In bats:

- Sensitive to high frequencies > 10000 Hz, and often well above the range of human hearing
- Echolocating bats use pulses of very high frequency vocal sounds (30 kHz to 120 kHz) in locating insect prey; a bat cry at 80,000 Hz has a wavelength of 4 mm, about the size of an insect
- A typical small bat has flexible pinnae about 5 cm long, and an ear canal of about 5 mm in length; resonant frequency range 17000 Hz down to 1545 Hz

Sound Perception: Human Sensitivity

A normal human ear has sensitivity to both intensity and frequency:

- The ear can detect extremely small differences in intensity in sound waves over a huge range: 1 x 10^{-12} W/m^2 to 1 W/m^2.
- The ear can distinguish different frequencies from 20 Hz to 20,000 Hz, though this range declines somewhat with age.

Individual ability to detect sound and discriminate between frequencies varies. Our ability to tell the difference between frequencies is called pitch discrimination. The human ear is optimized for detecting sound and discriminating frequencies in the range of the human voice, roughly 250-5000 Hz. The threshold of hearing is taken as 1 x 10^{-12} W/m^2 for a frequency of 1000 Hz; a much greater intensity would be required to hear frequencies below 200 Hz or above 10,000 Hz.

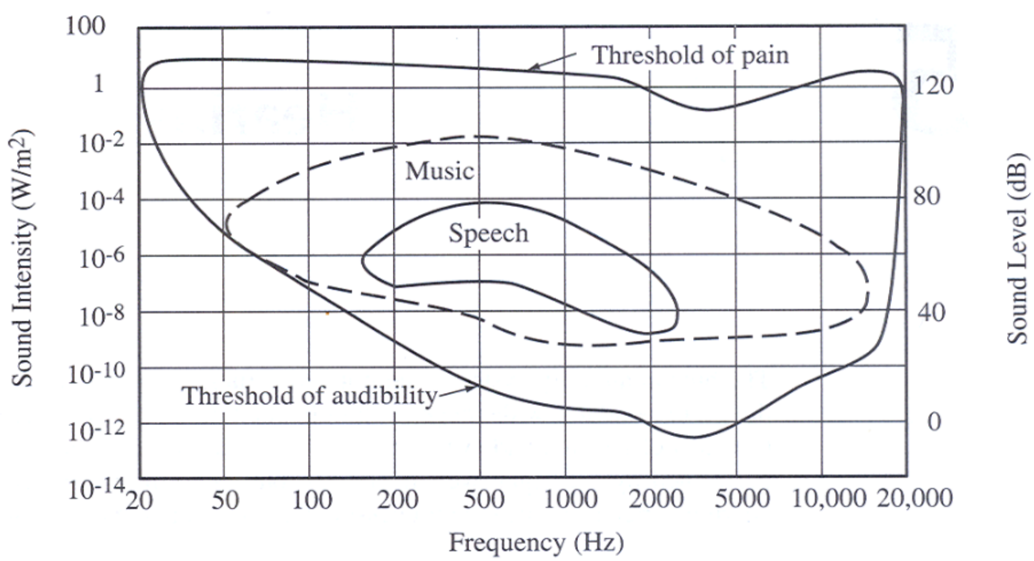


Figure 8. The Range of Human Hearing: Sound Intensity, Sound Level and Frequency. (Illustration adapted from Cameron *et al*.)

The physical term for sound intensity measured in Watts/m^2 is shown here on the left, while the related perceptual measure of loudness, sound level in decibels (dB) is on the right. The intensity or loudness detectable by the human ear is not the same for all frequencies – for example, the quietist sounds near 10 dB are not audible at low frequencies under 500 Hz, and the greatest sensitivity is in the middle range of intensities for frequencies in the human vocal range.

Sound Production: Evolution of Vocalization

The vocal cords are two skin-covered flaps of muscle that open and close across the passage to the lungs and oscillate with a wave motion as air rushes out during exhalation. These flaps first appeared about 300 million years ago in the colonization of land by animals that breathed with lungs rather than through gills. Initially the purpose of these folds of muscle was to act as natural valves, blocking everything but air from entering the respiratory system and allowing the ejection of phlegm from the lungs. Eventually, air-breathing animals with lungs discovered they could sound warnings or attract mates by blowing air out and intentionally vibrating the vocal cords. This kind of development is an illustration of what is known as Darwinian preadaptation.

Human Vocal Characteristics

Fundamental frequencies of human voices are from around 80 Hz to 1100 Hz. The frequency of the voice (or pitch) is the number of times the vocal cords vibrate per second (Hz). A bass singing a low A vibrates his cords 55 times/second; a soprano singing high C, vibrates hers at 1,047 times/second. The volume (amplitude, or loudness) of the voice is controlled by the magnitude of displacement of the vibrating object, i.e. the vocal cords.

So why do we sing?

The ancient natural rhythms of the earth – day length, lunar cycle, tides – instilled in all earth's organisms from bacteria to humans an equivalent sensitivity to rhythm and pattern. As Darwin proposed, pleasure seeking is an adaptive response across evolution – a sense of pleasure reinforces behaviours conducive to survival of a species (Darwin, On the Origin of Species by Means of Natural Selection, 1859); (Darwin, The Expression of the Emotions in Man and Animals, 1872) (Darwin, The Descent of Man, and Selection in Relation to Sex, 1871):

- Rhythm is pleasurable, and vocalization is necessary for communication, mate identification, warning, and social behaviours.
- Production of rhythmic vocalizations reduces stress and lowers blood pressure – singing is good for you!

The production of rhythmic or patterned communication is pleasurable for a number of adaptive reasons. Singing, especially in highly social animals, is a pleasure both in its choral and individual forms.

While listening to music is a great pleasure for many, it is (in my opinion) a much greater pleasure to make music, and to sing.

References

Allin, E. (1975). Evolution of the mammalian middle ear. *J. Morphol. , 147* (4), 403-437.

Cameron, J. R., Skofronick, J. G., & Grant, R. M. (1999). *Physics of the Body.* Madison, Wisconsin: Medical Physics Publishing.

Carroll, R. L. (1988). *Vertebrate Paleontology and Evolution.* New York: W.H. Freeman and Compant.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection.* London, UK: John Murray.

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex.* London, UK: John Murray.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals.* London, UK: John Murray.

Dorling Kindersley. (2012). *Prehistoric Life.* London: Dorling Kindersley Limited.

Hagen, E. H., & Hammerstein, P. (2009). Did Neanderthals and other early humans sing? Seeking the biological roots of music in the territorial advertisements of primates, lions, hyenas, and wolves. *Musicae Scientiae , 291* (13).

Hershberger, W. (2015). Biology of Insect Song. *The Songs of Insects* .

Manley, G. A., & Gleich, O. (1989). Evolution and Specialization in the Avian Auditory Periphery. In D. B. Webster, R. R. Fay, A. N. Popper, G. A. Manley, & O. Gleich (Eds.), *The Evolutionary Biology of Hearing* (pp. 561-580). New York: Springer-Verlag.

Searcy, W. A., & Andersson, M. (1986). Sexual Selection and the Evolution of Song. *17*, 507-533.