

Review

## The Variety of Information Transfer in Animal Sonic Communication: Review from a Physics Perspective

Neville H. Fletcher

Research School of Physics and Engineering, Australian National University, Canberra ACT 0200, Australia; E-Mail: neville.fletcher@anu.edu.au; Tel.: +61-2-6125-4406; Fax: +61-2-6125-0511

Received: 13 October 2009 / Accepted: 17 November 2009 / Published: 18 November 2009

---

**Abstract:** For many anatomical and physical reasons animals of different genera use widely different communication strategies. While some are chemical or visual, the most common involve sound or vibration and these signals can carry a large amount of information over long distances. The acoustic signal varies greatly from one genus to another depending upon animal size, anatomy, physiology, and habitat, as also does the way in which information is encoded in the signal, but some general principles can be elucidated showing the possibilities and limitations for information transfer. Cases discussed range from insects through song birds to humans.

**Keywords:** anatomy; physiology; acoustics; information; communication; environment

**Classification:** PACS 43.80.Ka, 43.80.Lb, 43.70.Bk

**Classification:** MSC 94A05, 94A17, 92-01

---

### 1. Introduction

Animals use a wide variety of techniques to communicate with other members of the same species, and occasionally with other predatory species. Some may be simply chemical, such as the use by insects of pheromones to lay food trails or by other animals to signal sexual availability or emotional state [1]. Similar information can be conveyed by visual signals such as the display of vivid plumage as a sexual attraction in birds, or the temporally patterned flashing luminescence of fireflies [2], but visual signals allow transfer of much more complex information sets ranging up to the encoding of information by

humans in drawings or in written texts. Of most importance for conspecific communication, however, is the use of acoustic signals, which have the advantages of spreading over a large area near the source and at the same time encoding a great deal of information in a given time. The way in which evolution has influenced the development of animal communication strategies is discussed in detail in a book by Hauser [3], while those seeking a detailed survey of all varieties of animal communication will find this in the comprehensive book by Bradbury and Vehrencamp [4] which gives many biological examples but with adequate physical and mathematical detail and copious references to the literature. The nature and purpose of animal signals has been critically examined in publications by Maynard Smith and Harper [5, 6], who highlight the compromises that must be made between the cost of signal production, its clarity, and its reliability. When we consider acoustic communication, general reviews have been given in books by Stebbins [7] and by Lewis [8], and there are many more specialised publications, some of which will be referred to later. The aim of the present paper is to review the field of animal auditory communication from a physics perspective, so that only a small number of specific cases will be examined, some of these representing extremes of what is possible. It will also not be possible here to enter into the quasi-psychological aspects of the subject, though these are indeed of great importance in real life.

The variety of auditory communication in animals is very great and the frequency range extends from below 8 Hz to above 100 kHz in different species. Frequencies above about 20 kHz are generally used for sonar exploration of the environment rather than for communication, but this frequency range is also used for close communication by some animals such as rodents. We will be concerned here mainly with communication in the atmosphere, but similar features are found in underwater communication. Mention should also be made of vibrational communication, which has many features in common with auditory communication, though there are some major differences.

From a physical or anatomical viewpoint animals can be broadly divided into two classes on the basis of anatomical and physiological differences. The first class is that of animals that do not have active respiratory systems driving their vocal apparatus. Prime examples are insects and crustaceans which generate sound signals by simple mechanical vibrations driven directly by muscles. The second class is the more familiar one in which air stored under pressure in a lung or respiratory sac is released through a vocal valve that it drives into oscillation to produce the sound. Birds, dogs and humans are familiar examples, but the class also contains sea animals such as dolphins and whales. A detailed exposition of the physical science underlying sound production and detection by these two classes of animals has been given in a book by the present author [9], and specific cases will be referred to later in the present review.

Acoustic communication has developed under the influence of evolutionary and environmental pressures but its objectives can be divided into two main classes. The first class aims to communicate over as large a distance as possible, as for example in shouting or loud speech by humans, while the second class is for confidential communication over short distances while minimising the spread of the information to other members of the same species or to predators, an example being whispering in humans. Much of the communication between animals of the same species will, however, lie in the space between these extremes, for example the desire to communicate over a distance large enough to reach all members of the local tribal group but using the smallest possible amount of physical exertion.

Acoustic communication strategies also involve other compromises relating to information content. Sometimes the message to be conveyed is extremely simple, such as a territorial marker or an alarm call, and has only small information content. The signal is then repeated frequently using a technique that ensures propagation to a maximum distance. At the other extreme, the signal may encode a high information content but require its audibility over only a relatively short range, an example being some types of formal human speech. Not all animals are able to make this distinction, but simple objective techniques have been developed by the more sophisticated species, as is discussed later. This subject has been discussed in some detail by Bradbury and Vehrencamp [4] and by Maynard Smith and Harper [6]. The whole topic is immensely complex and its understanding involves consideration of evolution, physiology, psychology, linguistics, and other similar disciplines. Only the simple physical aspects will be examined here.

## 2. Animal Size and Call Frequency

It is subjectively clear that the dominant frequency of the call produced by an animal is inversely correlated to the animal size. A scaling rule based upon simple classical mechanics suggests that, if all that changes is the size of the animal, then the call frequency should vary inversely with linear dimension or equivalently as the animal mass to the power  $-1/3$ . This is an oversimplification, of course, because physics also requires that the shape of the animal body must change with its size to preserve mechanical strength, but the prediction is remarkably close to what is actually observed [10], though there are some refinements to be included. Details vary, of course, between insects and mammals and between land and sea dwelling animals and it is also important to realise that this is only a general principle that is applicable on a large scale. Within a single animal species there is not necessarily any correlation between animal mass and vocalisation frequency—human sopranos are not always smaller in size than contraltos!

There is, in addition, a very large variation in sound production for various species when measured as acoustic power per unit body mass. Some species of Australian cicadas, such as the “Green Grocer” or the “Yellow Monday” (*Cyclochila australasiae*) are rated as the loudest insects in the world and produce a sound level in excess of 80 dB at 1 m, which is about 1 mW of sound power. Since the cicada has a mass of only about 1 g, this is equivalent to an output of about 1 W/kg. For comparison, a human shout is also about 80 dB at 1 m, which is only about 10  $\mu$ W/kg, or 100,000 times less per unit mass than that of the cicada. The reason is the biological importance of the signal, since a male cicada lives for only about 2 weeks after leaving its burrow in the ground and its only real purpose during that time is to attract a female. There is also a great difference in information content, as is discussed later.

Another interesting measure is the actual efficiency of sound production. For a typical mammal this is quite easily estimated. Taking the case of humans, the air pressure in the lungs during loud vocalisation is typically about 5 cm of water pressure, or 500 Pa, and the air flow is typically about 0.2 L/s or 0.0002 m<sup>3</sup>/s, which gives an input power to the vocal system of order 100 mW. The intensity of the sound produced is around 80 dB at 1 metre, which is equivalent to about 1 mW, so that the sound production efficiency is only about 1%, the remainder of the input energy being dissipated in viscous and mechanical losses. Interestingly, this is comparable with the efficiency of most musical wind instruments. This calculation, however, omits consideration of the efficiency with which food

energy is converted to muscle tension and motion, which further reduces the overall efficiency. Resonant systems, however, such as the tymbal of a cicada, can be much more efficient in producing sound with a fixed frequency in much the same way as a bell.

### 2.1. *Insects and crustaceans*

Insects and crustaceans generate a sound signal by exciting a part of their anatomy into vibration using muscular effort. An excellent overview is given in the book by Ewing [11] and in the collection edited by Kalmring and Elsner [12]. For sea-based animals the surrounding water is well matched in density to the vibrating element, so that the energy is rapidly radiated and the sound is almost a “click” which is then repeated regularly by muscular effort. For air-based insects the coupling between the vibrating element and the air is much less, so that its vibration continues for some time at its natural frequency. In many insects, such as crickets and grasshoppers, a wing or other thin panel is the vibrating element and it is excited by drawing across its edge a leg upon which is a regular array of ridges or saw-like teeth. Each ridge passage excites a burst of vibration in the panel but these are so close together in time that the sound seems almost continuous until the whole length of the leg has been traversed, when there is a small time interval until the excitation cycle repeats. Because both sides of the vibrating element are exposed in many cases, we have a dipole source which is not an efficient radiator since its two sides are in opposite phase. This dipole effect is reduced if one side of the panel is partially shielded, but some crickets actually make use of this dipole nature by locating it near a critical resonance position in a trumpet-like burrow that they dig in the soil [13]. This gives a much louder sound with power concentrated near the resonance frequency of the burrow. With all these systems the sound frequency scales about inversely with linear dimensions or with the cube root of the body mass. Typical frequencies lie in the range 3–5 kHz.

Other insects, such as cicadas, have rather different anatomy and an organ specifically evolved for efficient sound radiation. This consists of two flexible ridged diaphragms, or tymbals, which cover openings leading to a body cavity. Together these tymbals and the cavity constitute a resonant system with a well defined vibration frequency. This resonance is excited by contraction of muscles linked to the middle of each tymbal. Because the tymbals are ridged, they move inwards in sharp steps rather than smoothly, and each step excites the natural resonances of the tymbal membrane and cavity. Since the two tymbals are acoustically coupled and move in phase to constitute a resonant monopole source, this makes the radiation very efficient, giving the notably loud sound of the cicada species. Most cicadas have song frequencies in the 2–5 kHz range, and again varying about inversely with linear dimension. There are, of course, many anomalous cases in insects, a notable example being the bladder cicada *Cystosoma Saundersii* (Westwood) which is only about 5 cm long but most of this length is an abdominal sac of sufficient volume to reduce the song frequency to about 400 Hz [14].

This wide anatomical variety among insects means that simple scaling laws have many outliers, but the general rule of song frequency varying inversely with linear size is still broadly applicable, and the conspecific communication distance for insects is predicted by the theory to vary about as  $(\text{mass})^{0.5}$  [10].

## 2.2. Air-breathing animals

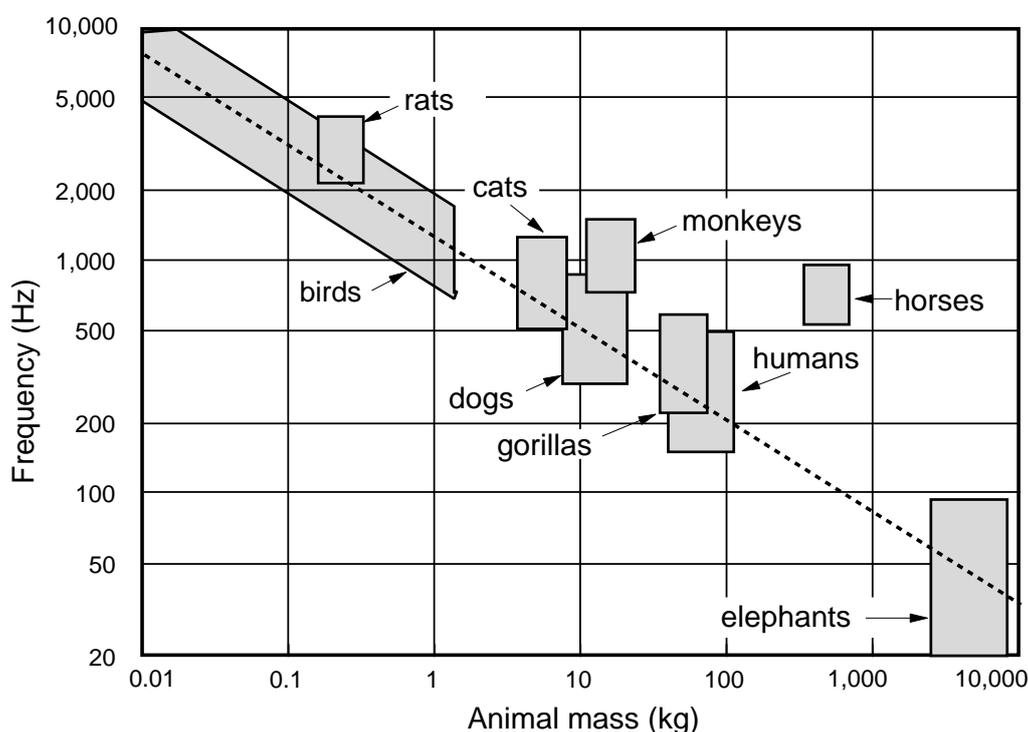
Air-breathing animals living on land produce sound by exhaling air stored under pressure in the lungs through a vocal valve, the larynx, that can be set into oscillation by the air flow, thus producing a periodically varying air flow out through the mouth. In most animals there is a single vocal valve located in the trachea, but songbirds instead have a pair of valves, one in each of the two bronchi, in an organ called the syrinx which is located just below the junction between the bronchi and the trachea. The air pressure in the lungs of any animal varies as the muscle tension multiplied by the muscle thickness and divided by the lung diameter, so that it should have nearly the same range independent of animal size, and this is typically around 200–1,000 Pa or 2–10 cm water-gauge, varying within this range with the type of vocalisation. A careful consideration of sound production, propagation attenuation, and hearing sensitivity leads to the conclusion that the optimal vocalisation frequency to achieve maximum conspecific communication range varies about as animal mass to the power  $-0.4$ , which is just a small correction to the power  $-0.33$  based upon simple linear scaling [10]. This power law is found to apply with moderate accuracy over the whole size range from mice to elephants, as shown in Figure 1, which has been modified slightly from the original version. The theory predicts that, assuming all animals to put equal effort into vocalisation, which is far from being generally true, the conspecific communication range should vary about as  $(\text{mass})^{0.6}$ .

Of course animals may not always aim to have conspecific communication at maximum distance, for they may often be traveling or living in a group. Their aim is then largely to communicate within the group at maximum efficiency so that they use the smallest possible amount of stored body energy. The optimal frequency for this is again about the same as that for maximum propagation distance. A third consideration is the presence of environmental noise, which may be due to wind, waves, or the calls of other animals. We return to this in Section 4, in connection with information transfer, but if the noise level is high then this may be a major factor in determining communication frequency. The most common form of natural noise is that in which the sound energy per octave band is about constant across the audible spectrum, so called “ $1/f$ ” or “pink” noise. The higher level of noise at low frequencies encourages animals to vocalise at rather higher frequencies than in a quiet environment, but this must be balanced against the higher level of atmospheric attenuation at high frequencies. The result is an optimal frequency that depends upon vocal power, and thus generally animal size, and also upon ambient noise level [9, 10], the expected scaling law being about  $(\text{mass})^{-0.27}$  [10], which is not very different from the quiet environment scaling law of  $(\text{mass})^{-0.4}$ . As a confirmation of this effect, some birds have been observed to sing at a higher frequency in a noisy city environment than the same species in their normal quieter rural habitat [15], perhaps for this reason.

When compared with insects, air breathing animals have very much greater flexibility in their vocalisations. The pulsating airflow through the larynx or syrinx contains large amplitudes of all the harmonics of the fundamental oscillation frequency of the valve and this air flow must pass through the upper vocal tract before being radiated from the mouth or beak of the animal. This upper vocal tract has many acoustic resonances that modify the relative amplitudes of these harmonics. If this vocal tract were a simple cylindrical tube, then these resonances would be at frequencies for which the tube length was an odd number of quarter-wavelengths, typically at about 500, 1500, 2500, ... Hz in the case of

humans. The radiated signal then has maxima, termed formants, in its spectrum near these frequencies. The vocal tract is, however, not a simple cylinder but can be modified greatly in shape by moving the tongue, jaws and lips, or alternatively the tongue and beak in the case of a bird. These shape changes shift the resonance frequencies of the tract by large amounts and this changes the spectral envelope of the radiated sound, giving the spectral peaks or formant characteristics distinguishing vowels in human speech. There are also common sounds in which the vocal valve does not vibrate but simply releases a short puff of air to produce a broad-frequency transient sound known in human speech as a consonant. These modifications to the simple output from the vocal valve are most important in the encoding of information, as will be discussed later.

**Figure 1.** Dominant frequency range of animal vocalisations as a function of body mass. The regression line shows  $(\text{mass})^{-0.4}$  as predicted by theory [10].



### 2.3. Aquatic animals

Aquatic animals will receive only brief mention in this paper, but a comprehensive treatment is given in a recent book by Au and Hastings [16]. It is helpful, however, to note the factors that are different from land-living animals. The first of these is the fact that the density of water is very close to that of animal flesh so that sound propagates through each of these rather than being reflected as at the density mismatch between animal flesh and air. This has led to the evolution of sensory detectors such as otoliths, which consist of a much denser small stone-like object supported on a hair which is connected to a neural transduction cell. Under the influence of a sound wave the otolith moves less than the supporting structure so that the hair cell is activated. For sound production the animals must also rely upon an air

motion through a vibrating valve between reservoirs to cause wall vibration, or else a modulated jet of air released into the water through the mouth or nostrils.

The other significant things are that sound absorption in water is much less than in air, so that there can be propagation over very long distances. This is also facilitated by the fact that the water depth in the ocean is limited so that spreading of the sound is effectively two-dimensional once the distance exceeds the ocean depth. In addition, the ocean often has a layered temperature or salinity profile, which also helps to make propagation two-dimensional. This results in a decrease in sound power as  $(\text{distance})^{-1}$  rather than as  $(\text{distance})^{-2}$  as in three-dimensional spreading.

Because of all these factors, the frequencies used for communication by aquatic mammals are very much higher than those used by land-dwelling animals of comparable size. A large whale, for example, may generate calls with fundamental frequency above 1 kHz while an elephant of comparable mass uses a frequency well below 100 Hz. In addition, animals such as dolphins generate whistles in the range 15–30 kHz and echo-location calls with frequencies in the range 50–100 kHz, which is comparable with the frequencies used by bats.

#### 2.4. Directionality

The sound signals produced by animals generally have directional characteristics and usually a single direction in which the sound level is greatest. This is of obvious advantage in communication with other individuals, but a disadvantage when communicating with a large group. For a simple sound source such as an open mouth, directionality becomes significant at wavelengths shorter than about twice the mouth diameter [9], or above about 4 kHz for a mouth of diameter 2 cm. This is obvious in human speech, for example, where the consonants that encode much of the meaning are difficult to hear when the speaker is facing away. To communicate with a surrounding group in large vertebrate species, therefore, the signaling animal must generally move its head around. In the case of insects, the sound source is generally smaller than the dominant wavelength so that the signal is nearly omni-directional.

Directionality is even more important in the animal receiving the sound signal, for it allows the location of the sender to be determined and also serves to maximise the signal strength against environmental noise. While a single large diaphragm with a neural transducer would provide some directionality, much more efficient auditory systems have evolved that use a pair of small diaphragms coupled acoustically through a cavity or a tube, each diaphragm being connected to a neural transducer. Such a system can provide good directional sensitivity at a particular frequency [9] which is generally the dominant frequency for conspecific communication. In mammals these individual diaphragms are generally augmented by the presence of external horn-shaped ears which amplify the acoustic signal and also provide additional directivity [18] and this system has evolved further by replacing the acoustic coupling between the two ears with neural coupling through a part of the brain and presumably neural evaluation of the phase and amplitude difference between the two signals rather than simple detection of the result of their acoustic interference.

### 3. The Variety of Call Types

The amount of information contained within a sonic signal varies widely with the genus of the animal involved and the evolutionary function of the signal [3, 6]. A careful comparison of two related mammalian species has been given by McCowan *et al.* [17] but here we examine the wide variety of acoustic communication strategies involved by examining a few very different animal genera.

#### 3.1. Cicadas, crickets and other insects

While the variety of insects and their communication methods is very great [11, 12], the simplest signals are probably those of crickets or cicadas, the purpose of which is to advertise presence and location so as to attract females for mating. These signals generally have a well defined and stable dominant frequency and spectral envelope as determined by the anatomy of the signaler and the repetition rate of the muscle contraction or leg motion exciting the resonator. For a typical cicada the sound spectrum has a dominant frequency close to 3 kHz, a frequency spread of about  $\pm 1$  kHz, and a pulse repetition rate of about 2 per second. All the cicadas in a given area emerge from the ground at about the same time in the summer, sometimes after 6 or 7 years underground as nymphs in the case of some Australian cicadas or as long as 13 or 17 years for some North American cicadas, so that trees in an area are very fully populated by singing males for a few weeks in summer. The combined chorus generally has random muscle-contraction timing, but occasionally the whole group of more than a dozen insects can lock into synchrony for as long as 10 seconds. The evolutionary purpose of the song is to attract females, first to the group as a whole and then to individual members. The only information contained in the song appears to be the species of the insect, as encoded in the frequency and repetition rate, the location of the group, as particularly emphasised in the synchronised intervals, and later at short range the location of individuals within the group. The fact that the individual songs are broken into repetitive segments certainly helps with this last objective. The other item of information content, which is common to the songs of most animals, is to indicate vitality so as to attract a mate, a piece of information that is probably encoded in the acoustic intensity of the song.

The songs of crickets are similar to those of cicada, though the emphasised frequency may be higher because of the smaller size of the insect. The cricket burrow helps to increase the radiated sound intensity but crickets tend to sing as isolated individuals rather than as large flocks so there is little nearby competition. The information content is again essentially just species, location and vitality.

#### 3.2. Birds

Birds are perhaps the class of animals that has the most diverse range of vocalisations. While there is a fairly good correlation between dominant song frequency and body mass [10, 19], there is quite a wide scatter and a great variety in the structure of the songs.

As noted before, there are two classes of avian vocal anatomy, one with a single vibrating vocal valve in the trachea, very much as in mammals, and one with a pair of valves in the bronchial tubes below their junction with the trachea. This anatomical variation does not define the song classes, but songbirds with dual syringeal valves generally have a wider repertoire and even the ability to sing two notes at the same time. There have been many publications in the biological literature on this subject, a good collection

being that edited by Kroodsma and Miller [20], and also publications on the underlying acoustics [21]. The avian auditory system is similar to that in reptiles and essentially consists of a tube joining the two tympana with the physical parameters adjusted to give optimal sensitivity and directionality close to the conspecific song frequency [9] and covers a frequency range of about 20 Hz to 10 kHz with maximum sensitivity in the range 1–3 kHz, which is similar to that of humans.

There are, however, many variations to this generic system. Birds such as ravens produce calls with a strong harmonic content which is filtered by the upper vocal tract and beak to give emphasised formant bands much as in human speech, though at a higher frequency. The information content of these calls can therefore be large because of the number of parameters involved. At the other end of the complexity scale we find the “coo” of doves, which is a nearly pure-tone call lasting for only about a second, the underlying mechanism being an inflatable sac into which the bird sings with its beak closed, the sound being radiated by the vibrating sac walls [22]. Intermediate between these two are the pure-tone calls of birds such as the Northern Cardinal which are sung with an open beak and an adjustable vocal sac which is tuned to follow the song frequency [23] which may sweep by more than a factor 2 in frequency in about a second. Finally, mention should be made of the class of chaotic “shriek” calls made by many species of cockatoos, but particularly the large Australian sulphur-crested cockatoo *Cacatua galerita* [24]. The information content here is low but the purpose of the call appears to be to define territory on behalf of a considerable flock of these birds. These cockatoo calls have a maximum in the spectrum between 2 and 3 kHz so that they sound very loud to humans. The birds themselves are very beautiful but delight in tearing flowers and new branches off trees and rubber gaskets off street-lights.

Some cockatoos and parrots can be taught to imitate human speech, learning phrases such as “Pretty Polly”. The imitation is quite intelligible but lacks emphasis on the lower frequencies so that it sounds like a telephone of poor quality. There is, however, another Australian bird, the Superb Lyrebird *Menura novaehollandiae* [25], which has carried mimicry to a supreme level. Not only can it imitate quite faithfully the calls of other birds, even those as different from it as the laughing kookaburra, but it also produces convincing versions of mechanical sounds such as motor exhausts and chain-saws. This vocal ability, along with its beautifully spectacular tail feathers, is presumably intended both to define territory and to attract a mate. While the amount of information potentially encoded in the signal is very large, it probably conveys only a limited and qualitative amount to the conspecific listeners.

### 3.3. Mammals

Mammals range in size from mice to elephants and their dominant vocalisation frequencies from over 3 kHz down to as low as 20 Hz. Because they generally live in social communities and have mental capacities greater than other animal types, their vocal signals have become very sophisticated and largely designed to convey specific information to other members of the community. Large animals such as elephants can be heard by other elephants at distances up to about 10 km in the evening quiet period, but at that distance only the fundamental is audible and this conveys no information other than existence. In the case of smaller animals the purpose of conspecific communication is to warn of predators, locate food supplies, guide young offspring, and seek mates. Some animals, such as dogs, also use loud abrupt “barking” signals to warn off potential intruders. A wide-ranging survey of mammalian vocalisation is provided in a recent book edited by Brudzynski [26].

It is interesting that some mammalian species have developed two types of communication strategy suited respectively to broad communication over moderately large distances and more confidential communication with family members over short distances. In the case of small rodents with a normal vocalisation frequency in the range 3–5 kHz, the short-range communication is often carried out in the frequency range 20–30 kHz, obviously using a different sound-production strategy [27]. In the more familiar case of humans the close-communication strategy is termed “whispering” and involves using a broadband turbulent noise signal produced by air flow through a fixed aperture or over a sharp edge in the mouth and then shaping the spectrum of this signal to recognisable vowels through tuning the formant frequencies by adjustment of the jaw, tongue and lips.

Since human vocalisation is familiar to most people and well documented [28], it will not be explored in further detail here, the brief summary in Section 2.2 being sufficient. There is, however, a great difference between information coding in human speech and in the sounds made by other mammals. Instead the discussion in the following sections will concentrate on the means and effectiveness of information transfer, largely by humans though the results apply at least qualitatively to other animals that have the equivalent of speech. More restricted cases will also be considered.

#### 3.4. *Vibrational communication*

As well as auditory transmission of signals, most animals have at least some sensitivity to vibration of objects in which they are in contact, and some insects have developed specialised “sub-genual organs” in their leg joints to detect vibration of the substrate upon which they are standing. A comprehensive review of the subject has been given by Hill [29]. There are two broad classes of transmission channels for vibrations, the first being simply the ground, with the vibration propagating as a circularly spreading surface wave and thus broadcasting the signal, usually made by foot stamping as in elephants, to all animals standing on the ground. The second applies to animals such as insects that live on trees or smaller plants. In this case the signal propagates as a bending wave or a shear wave depending upon frequency, but only along a branch of the tree or plant, and there are strong reflections at any junctions. Transmission is therefore generally limited to other insects on the same branch. In all these cases the conspecific information transfer rate is much less than for vocalisation because there is no specialised organ for generating the vibrational signal.

A particularly interesting form of vibrational information transfer is that used in the “waggle dance” of honeybees *Apis mellifera*, which signals the location and distance of food sources and other items of interest by repetitive dance motion in a figure-eight pattern as described by Frisch [30]. More recent studies, as cited by Nieh and Tautz [31], show that the transfer of vibrational information about the geometry of the dance is a complex process, and there is even the possibility that the existence of subgenual vibration detectors on each of the six legs of the bees may be important.

Another distinctive case is that of insects that scavenge other smaller insects that have been trapped by surface tension on the surface of a pond of still water. The scavengers have long legs and hydrophobic surfaces on their feet so that they can stand safely on the water surface, supported by surface tension forces. They can then detect the slowly propagating surface waves generated by the trapped insect as it struggles to free itself from the water. In this case there is no intention on the part of the trapped insect to transmit information, but this happens anyway and reveals its location from the propagation direction

of the waves, and also something about its size from the frequencies involved. The same is true for a spider detecting smaller insects trapped in its web.

#### 4. Information Transfer

Animals use sonic signals to transfer a wide variety of information types, and there is great variation between the information that could be transferred on a given signal and the information that is actually transferred, so both of these aspects must be examined. A major distinction is the way in which information is encoded in the sound signal. We are most familiar with human speech in which information is transferred using words, which are made up of syllables, which are in turn made up of phonemes or elementary vocal units. This is a very flexible system and allows transfer of a very large variety of information. Most other animals, however, do not compose their vocalisations from such identifiable small elements but use longer passages of sound to convey one of a limited number of messages. Since the human language system is the more efficient for diverse information transfer, most of the following discussion will be based upon it, but the same principles apply to other encodings.

There is, however, a disadvantage associated with the complex structure of phrases and sentences in human speech and this is the possibility of deception and dishonesty [32]. This topic is too complex for consideration here but is the subject of much contemporary research. At a higher level the question of the origin of animal communication might also be examined in the light of what benefits it brings to the animals involved. This question, which is again outside the scope of the present paper, has been considered in terms of computer game simulation by Tanimoto [33].

The relation between information and entropy was first clearly defined by Shannon [34], who also considered the three major elements of the information chain: the source, the transmission channel, and the receiver. While the source determines the method of encoding the information and the actual information to be placed upon the signal, the transmission channel generally degrades this information by attenuation and by the addition of noise. The receiver must then retrieve as much as it can of the original signal and finally convert it to information. While the basic transmission channel—sound propagation through the atmosphere or ocean—is quite well understood, there is the complication of added noise which depends upon other activities in the environment. Sound reception by animals is also quite well understood, but the final phase in which the received signal is decoded into information relevant to the receiving animal is essentially psychophysical and varies greatly from one genus to another.

A major reason for this large variation is evolutionary and relates to the purpose for which the signal is used, and this depends on population density, environment, existence of major predators, and life-span of individuals. This may lead to huge variations between major genera in energy expended on signaling and also to sophisticated differences in signal coding for advanced animals. An excellent review of this matter has been given in Part II of the book by Bradbury and Vehrencamp [4], but here we will be able to summarize only the basic aspects of information transmission.

Information is most conveniently measured in binary digits or “bits”, each bit being either 0 or 1. This concept was perhaps introduced originally in Morse Code signals but is now familiar in the world of computing where text, sound and visual pattern can all be digitally encoded in bits. The information

content of a continuous signal of acoustic power  $P$  and frequency bandwidth  $W$  (hertz) propagating in a channel that adds noise power  $N$  was shown by Shannon [34] to have a maximum value of

$$C = W \log_2 \left( 1 + \frac{P}{N} \right) \quad (1)$$

where the logarithm base is taken as 2 so that  $C$  is measured in bits per second (bps). Animal communication, however, does not even approach this maximum value because of the limitations imposed on the encoding and decoding systems by the anatomical and neural structures employed. In the case of no added noise  $C \rightarrow \infty$  and, even for a signal-to-noise ratio of 1:1,  $C \rightarrow W$  which implies about 3,000 bps for a channel of width 3 kHz as is typical of human speech, a figure that is vastly in excess of what can actually be achieved. Part of the underlying reason for this shortfall is that animals mostly encode information in a discrete manner with a limited number of elements such as the words in a short dictionary, and they have limited time-resolution for both encoding and decoding, as we now examine.

The sonic signal generated by an animal has its potential information content encoded in the time-varying spectrum  $X(f, t)$  of the sound. To quantify the information content the time variable  $t$  is taken to be divided into successive elements of length  $\Delta t$ , where the smallest value of this time interval is the time response limit of the articulatory vocal system of the sender and of the auditory system of the receiving animal, typically about 0.05 seconds. The frequency scale is not divided in such a simple way because it varies greatly with the animal and with the information encoded. The three vocal formants in human speech [36], for example, basically encode just 5 vowels in short form plus another 5 long-form versions and about the same number of diphthongs, making about 16 elements in all in non-tonal languages. The count is rather different for tonal languages such as Chinese, but the final number of vowel elements is not much different. In contrast, a human singer with a range of two octaves must encode the pitch of 24 semitones with an accuracy of better than half a semitone and often with vowels superimposed, making about 64 elements in total. In the case of human speech, about 16 consonants must also be included. The figures here have been approximated by powers of 2 so that a signal element with  $n$  possible forms is taken to encode  $\log_2 n$  bits. The encoded vowels in speech thus represent about 4 bits, and the attached consonants another 4 bits, making 8 bits in total. For an operatic singer the encoding might be as high as 12 bits when musical pitch is included, but the pitch encoding generally takes place at a much slower rate than does normal speech encoding, so that this will not be considered further here.

The information content of signal depends, however, not just upon the possible variety of its elements but also the rate at which these can be produced, received and decoded, and this rate is generally significantly slower than the response rate of the receiver. In addition, not all possible signal components are “allowed” in the sense that they carry information. In human speech or song as discussed above, for example, sequences of syllables are meaningful only if they constitute words. Shannon’s formal definition [34] of the capacity  $C$  of a discrete communication channel is

$$C = \lim_{T \rightarrow \infty} \frac{\log_2 n(T)}{T} \quad (2)$$

where  $n(T)$  is the number of “allowed” signals of duration  $T$ . In human speech, for example, the duration of a voiced syllable is about 0.2 seconds, with each syllable consisting of a vowel and a consonant and

thus containing about 8 bits of information. This makes a maximum information rate of about 40 bits per second (bps) if all syllable sequences are possible. For Morse code transmission, as discussed by Shannon [34], there are just 32, or  $2^5$ , allowed symbols and the transmission rate to a human listener is only about 3 symbols per second on average, so that the information rate is only about 15 bps.

When only “allowed” and thus meaningful speech signals are considered, however, we find that a typical dictionary of the English language contains about 10,000 words with lengths ranging from 1 to more than 5 syllables and with many of them having variants relating to their grammatical form. Most of these words, however, are not commonly used in conversation, the number being more like 1,000 for “academic” speech or as low as 100 for “common” speech (even by academics). The 1,000 word set typically has an average word-length of about 3 syllables while the “common” set is between 1 and 2 syllables, so that the values of  $T$  are about 0.7 s for 10,000 words, 0.3 s for 1,000 words, and 0.2 s for 100 “common” words. Applying Equation 2 to these figures gives channel capacities of about 20, 30 and 35 bps for these three cases. Perhaps surprisingly, the complete “allowed” set is least efficient in information transfer while the least extensive word set conveys information most rapidly.

Here, of course, we encounter the problem of the meaning of the word “information” and its difference from the semantic notion of “meaning”. If the length of a word is taken to be 2 to 3 syllables, then the number of possible 3-syllable combinations is of order  $10^7$  and the number of possible 2-syllable combinations is about  $3 \times 10^4$ , both numbers being much greater than the number  $10^4$  of words in the dictionary. On top of this, the individual words must usually be combined into sentences, or at least phrases. This means that when the sonic signal is designed to encode and transmit human speech there must be very strong restrictions placed upon the allowed signals if they are to have any meaning to the receiver. This situation was investigated by Shannon [34] for the case of written rather than spoken language and we will not pursue it further here because of its complexity.

At another level, ordinary “social” speech using simple words may really just convey warnings, location, or emotional feelings rather than “intellectual” information, while complex “academic” speech can be either rich in complex information or, like the famous Social hoax [35], completely meaningless. This is not the place to enter into such complexities but rather to examine the differences between possible acoustic signaling capacities for various animals. Detailed discussions of speech perception by humans have been given by Moore [37] and by Miller [38], and a much broader treatment for animals in general by Bradbury and Vehrencamp [4].

As an extreme version of simple communication let us consider the signals of a cicada. As discussed before, these consist of a repetitive sound with a frequency of about 3 kHz and a repetition rate of about 2 pulses per second. Apart from existence and location, this signal encodes simply species and perhaps vitality, the physical variables being just frequency, loudness and repetition rate. Since there are rarely more than three or four competing cicada species in any neighborhood, only this number of different frequency bands or pulse rates are required for differentiation, and the information is essentially “hard wired” into all members of a given species. The loudness signal for vitality similarly is automatic and probably has no more than two values from the viewpoint of the receiving female which simply aims to choose the loudest male. According to Equation 2 this gives an information content  $C \approx 3$  bps.

Note that spatial location does not appear in Shannon’s formula because it is information of a quite different type. A steady single-frequency signal could be used by a suitably endowed listener to locate

the source with arbitrary precision, and the same is true of a broad-band “click” pulse of very short duration. This directional information will therefore not be considered in the subsequent discussion.

An interesting example of the importance of transmission loss occurs in the case of the vocal calls of elephants. As reported before, the elephant call has a fundamental frequency as low as 20 Hz and a great deal of acoustic power so that it can be heard by other elephants over distances as large as 10 km. Information is encoded in the call by variation in the frequency envelope of the sound, which has many higher harmonics, but these higher harmonics are more rapidly attenuated by atmospheric absorption than is the fundamental, so at the limit of distance this is all that is heard. Behavioural studies show that other elephants are unable to recognise the individual who is the source of the sound unless harmonics as high as 100 Hz are audible, since the spectral envelope rather than the fundamental frequency encodes the necessary information.

#### 4.1. *Repetition, imitation and meaning*

A common feature of almost all vocal signals is that they are repetitive. An extreme example is the case of the cicada, in which the information is contained essentially in the frequency and amplitude of the signal pulse and this pulse is repeated thousands of times every hour. While this might be regarded as a very primitive behaviour, it could be said to be analogous to the concerts and recordings of rock-music singers! The aim is to ensure that the message encoded in the call—largely one of identity—is received and correctly interpreted.

Birds provide a good example of an intermediate state in which one of a small number of often complex songs is produced repeatedly and with little variation. Sometimes the objective of the repetition is to continually assert a territorial claim, but the songs may also represent calls to females and demonstrations of the attractiveness of the singer or messages about a food source. The Australian Superb Lyrebird, referred to before, which mimics the calls of other birds or the sounds it hears in the environment, aims simply to convey a signal of its ability and attractiveness in the same way as would a display of tail feathers, and any species-specific information encoded in the original call is repeated without reference to its meaning. The call may also contain what is effectively a “signature” so that it signals the identity of the caller to other member of its community.

On the whole, the vocalisation of mammals probably conveys more actual information to the hearer than for other species, which is not surprising given their scale of mental development. This applies particularly to primates, which approach human abilities, but is also probably true of whales, dolphins and other sea-living mammals. Some quite intelligent animals such as dogs, however, have not evolved to use sonic communication to anything like the same extent as others.

The conclusion to be drawn from this is that, while a mathematical analysis of the structure of a signal gives useful detail about the amount of information that is formally encoded within it, this formal definition of “information” differs from the psychophysical interpretation of the term, which involves the extent to which the sender of the message has encoded this information in the signal, either purposely or inadvertently, and the extent to which the receiver is able to decode the signal structure into information that is meaningful to it. A helpful discussion is given by Bradbury and Vehrenkamp [4].

#### 4.2. Noise

Because animals live in environments that are never truly silent and are often very noisy, it is important to examine the effect of this ambient noise on the transmission of information. This matter was treated in detail by Shannon [34] and has already been alluded to in Equation 1 but only a short discussion specific to animals will be given here. Suppose there is entropy  $H(x)$  in the signal injected into the transmission medium and entropy  $H(y)$  in the received signal. In the noiseless case the joint entropy of input and output will be  $H(x, y) = H(x) = H(y)$ , but in the general case where there is transmission noise we define “conditional entropies”  $H_x(y)$  and  $H_y(x)$ , the entropy of the output when the input is known and vice versa. We then have the relation

$$H(x, y) = H(x) + H_x(y) = H(y) + H_y(x) \quad (3)$$

The actual rate  $R$  of transmission of information is then

$$R = H(x) - H_y(x) \quad (4)$$

Shannon examines this relation in detail for the case of binary-encoded digital information, but similar conclusions are reached in more general cases. The received information content is less than the injected content and the difference increases with the level of interfering noise, ultimately becoming zero.

In order to maximise the reliability of information transmission, animals have evolved to have maximum hearing sensitivity over a limited frequency band that encompasses the conspecific vocalisation frequency, including the spread involved in the encoding. For humans, for example, hearing sensitivity is greatest over a band from about 500 Hz to 5 kHz, and this covers the frequency range of the first five formant resonances which encode the vowels, together with the broad band and transient response that encodes the consonants. The fact that the fundamental frequency of about 100–200 Hz for males is not included in this range does not matter greatly, since the brain can derive the fundamental frequency from the spacing of its harmonics, as in the ability to differentiate between male and female speakers on a telephone. Other animals generally have hearing abilities similarly tuned to match their vocalisation frequencies as discussed in Section 2..

The other interference-minimisation strategy in the case of environmental noise is to use as high a frequency as possible consistent with the aim of also maximising audibility distance. A third strategy with the objective of maximising information reliability rather than information content is to slow the encoding rate so that the essentially random environmental noise averages towards a constant value and  $H_y(x) \rightarrow 0$ . Cicadas, as described in Section 3.1., have a repetitive call with constant spectrum and a duration of order 0.5 s, so that it would be very difficult to misinterpret their presence even with a high ambient noise level. In human speech, competent orators slow their speech rate compared with conversation and exaggerate the differences between various vowels and consonants with the same objective.

Several studies have examined the influence of noise upon the communication strategies of various animals. Slabbekoorn and Peet [15] showed that some species of birds purposely sing at a higher pitch in a city environment than they do in quiet countryside, while Doyle *et al.* [39] found that Humpback whales increase the rate and repetitiveness of their communication calls under the influence of marine

vessel noise. This strategy is a variant of simply extending the duration of each component of the call and has the advantage that it avoids complete obscuration of a part of the call by a long burst of noise.

## 5. Music and Information

Birds and some other animals use variations in the frequency of their calls to encode certain types of information, but this technique is greatly expanded in the case of musical compositions and performances by humans. We set aside here the case of songs, which also include speech encoding, and consider only “pure” music such as produced by musical instruments. A musical composition, if we leave out of consideration modern computerised “sonifications” of natural sounds or patterns, then consists of a set of musical notes, each having fundamental frequency and temporal location as its prime attributes but with spectral envelope and loudness as subsidiary characteristics. A formal musical composition generally has this set of notes arranged into several sub-sets which overlap in time.

Setting aside the characteristics of spectral envelope and loudness, which generally serve mostly to distinguish one sub-set of the total note set from another, a composition then typically has up to about four or five note sub-sets that overlap in time, but generally with a high degree of time-correlation between them. In a “harmonic” composition this correlation is almost exact synchronisation, while in a “contrapuntal” composition such as a fugue there is also a high level of correlation but with a considerable time delay between sub-sets.

The encoding of musical information is essentially discrete rather than continuous, since frequencies are normally specified in semitones, with a frequency ratio  $2^{1/12}$ , over a frequency range of about 7 octaves or a factor  $2^7$ . This gives a total of about 84 possible pitches in normal Western music [40], though perhaps more realistically about 50 spread across 4 octaves, which is a little less than  $2^6$  or 6 bits per note. A realistic maximum for note sequences is about 4 per second, though faster sequences can be used over very short periods, so this gives an upper bound of about 24 bps per note sub-set. Given up to 5 note sub-sets yields the impressive value of about 120 bps. This is, however, an upper limit and no human listener could decode information at this rate for more than a few seconds. A more realistic value is perhaps that for a four-part fugue played at about 2 notes per second, which gives about 48 bps. Even this is an overestimate, since the pitch changes in a single note sub-set are generally nearly normally distributed around a central pitch and with a standard deviation of only about 5 semitones. Not many people could decode information even at this rate, and fortunately the occurrence of repetitive patterns in the composition significantly reduces the formal information content. Some notable musicians, however, are known to have been able to memorise such complex compositions with durations of several minutes in a single hearing and later to write them down.

## 6. Conclusions

Sonic communication in animals has evolved over millions of years with the primary objective of communication with other members of the same species. The information content of the communication varies with the mental development of the animal concerned, so that insects generally communicate only their identity and location, the next range adds territorial identification and alarm signals, while higher animals signal sexual attractiveness and other matters as well. Details of the vocalisation frequency used

and the audible range achieved depend upon the size of the animal in a broadly predictable way, though there are some outliers. Near the top of the evolutionary tree, primates and finally humans communicate much more subtle and abstract ideas as well. The fact that there are many quite different human languages, all constructed from similar sound elements such as vowels and consonants, plus whistles in some cases, shows that there is no really optimal encoding for information in human speech at higher levels, though its elements are constrained by the sound generation mechanism and tract resonances of the vocal system. Many animals, including humans, also use different communication strategies for long-distance and short-distance conspecific communication, with the objective of keeping their identity and information safe from predators or even from other members of the same species.

At the highest level, the information content of animal communication, and particularly of human speech, is almost impossible to gauge except with a primitive definition of “information”. In particular, since most vocal signals are rather repetitive, they may well be communicating information already known to the listener, so that the content of “new” information is small though the formal information content may be large, as in “boring” conversation! A discussion of these matters, however, would take the discourse out of the realm of biophysics into that of psychophysics, which is not part of the area explored here.

### Acknowledgments

As a physical scientist, I must confess to a lack of detailed knowledge of the biological literature, so that the citations in this article are very selective and many refer to collaborative projects in which I have been involved. I am most grateful to these colleagues. My purpose in the present paper has been to outline some of the physical principles underlying animal communication, and it is perhaps more appropriate that well-informed biologists examine the application of these principles to specific cases.

### References

1. Wyatt, T.D. *Pheromones and Animal Behaviour: Communication by Smell and Taste*; Cambridge University Press: Cambridge, UK, 2003.
2. Lloyd, J.E. Bioluminescence and communication in insects. *Ann. Rev. Entomol.* **1983**, *28*, 131–160.
3. Hauser, M.D. *The Evolution of Communication*; MIT Press: Cambridge, MA, USA, 1966.
4. Bradbury, J.W.; Vehrenkamp, S.L. *Principles of Animal Communication*; Sinauer Associates: Sunderland, MA, USA, 1998.
5. Maynard Smith, J.; Harper, D.G.C. Animal signals: models and terminology. *J. Theor. Biol.* **1995**, *177*, 305–311.
6. Maynard Smith, J.; Harper, D. *Animal Signals*; Oxford University Press: Oxford, UK, 2003.
7. Stebbins, W.C. *The Acoustic Sense of Animals*; Harvard University Press: Cambridge, MA, USA, 1983.
8. *Bioacoustics: A Comparative Approach*; Lewis, B., Ed.; Academic Press: London, UK, 1983.
9. Fletcher, N.H. *Acoustic Systems in Biology*; Oxford University Press: New York, NY, USA, 1992.
10. Fletcher, N.H. A simple frequency-scaling rule for animal communication. *J. Acoust. Soc. Am.* **2004**, *11*, 2334–2338.

11. Ewing, A.W. *Arthropod Bioacoustics: Neurobiology and Behaviour*; Cornell University Press: Ithaca, NY, USA, 1989.
12. *Acoustic and Vibrational Communication in Insects*; Kalmring, K., Elsner, N., Eds.; Verlag Paul Parey: Berlin, Germany, 1985.
13. Daws, A.G.; Bennet-Clark, H.C.; Fletcher, N.H. The mechanism of tuning of the mole cricket singing burrow. *Bioacoustics* **1996**, *7*, 81–117.
14. Fletcher, N.H.; Hill, K.G. Acoustics of sound production and of hearing in the bladder cicada *Cystosoma Saundersii* (Westwood). *J. Exp. Biol.* **1978**, *72*, 43–55.
15. Slabbekoorn, H.; Peet, M. Birds sing at higher pitch in urban noise. *Nature* **2003**, *424*, 267.
16. Au, W.W.L.; Hastings, M.C. *Principles of Marine Bioacoustics*; Springer: New York, NY, USA, 2008.
17. McCowan, B.; Doyle, L.R.; Hanser, S.F. Using information theory to assess the diversity, complexity, and development of communicative repertoires. *J. Comp. Psych.* **2002**, *116*, 166–172.
18. Fletcher, N.H.; Thwaites, S. Obliquely truncated simple horns: Idealized models for vertebrate pinnae. *Acustica* **1988**, *65*, 194–204.
19. Ryan, M.J.; Brenowitz, E.A. The role of body size, phylogeny, and ambient noise in the evolution of bird song. *Am. Nat.* **1985**, *126*, 87–100.
20. *Acoustic Communication in Birds*; Kroodsma, D.E., Miller, E.H., Eds.; Academic Press: New York, NY, USA, 1982.
21. Fletcher, N.H.; Tarnopolsky, A. Acoustics of the avian vocal tract. *J. Acoust. Soc. Am.* **1999**, *105*, 35–49.
22. Fletcher, N.H.; Riede, T.; Beckers, G.J.L.; Suthers, R.A. Vocal tract filtering and the “coo” of doves. *J. Acoust. Soc. Am.* **2004**, *116*, 3750–3756.
23. Riede, T.; Suthers, R.A.; Fletcher, N.H.; Blevins, W.E. Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc. Nat. Acad. Sci. USA* **2006**, *103*, 5543–5548.
24. Fletcher, N.H. A class of chaotic bird calls? *J. Acoust. Soc. Am.* **2000**, *108*, 821–826.
25. Smith, L.H. *The Lyrebird*; Lansdowne Press: Melbourne, Australia, 1968.
26. *Handbook of Mammalian Vocalization*; Brudzynski, S.M., Ed.; Academic Press: Oxford, UK, 2009.
27. Brudzynski, S.M.; Fletcher, N.H. Rat ultrasonic vocalization: short-range communication. In *Handbook of Mammalian Vocalization*; Brudzynski, S.M., Ed.; Academic Press: Oxford, 2009; pp. 69–76.
28. Lindblom, B.; Sundberg, J. The human voice in speech and singing. In *Springer Handbook of Acoustics*; Rossing, T.D., Ed.; Springer: New York, NY, USA, 2007; pp. 669–712.
29. Hill, P.S.M. *Vibrational Communication in Animals*; Harvard university Press: Cambridge, MA, USA, 2008.
30. von Frisch, K. *The Dance Language and Orientation of Bees*; Harvard University Press: Cambridge, MA, USA, 1967.
31. Nieh, J.C.; Tautz, J. Behaviour-locked signal analysis reveals weak 200–300 Hz comb vibrations during honeybee waggle dance. *J. Exp. Biol.* **2000**, *203*, 1573–1579.
32. Lachman, M.; Bergstrom, C.T. The disadvantage of combinatorial communication. *Proc. Roy. Soc. Lond. B* **2004**, *271*, 2337–2343.

33. Tanimoto, J. What initially brought about communications? *Biosystems* **2008**, *92*, 82–90.
34. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1962.
35. Sokal, A. *Beyond the Hoax: Science, Philosophy and Culture*; Oxford University Press: Oxford, UK, 2008.
36. Zwicker, E.; Fastl, H. *Psychoacoustics: Facts and Models*; Springer-Verlag: Berlin, Germany, 1999.
37. Moore, B.C.J. *Introduction to the Psychology of Hearing*; MacMillan: London, UK, 1977; Chapter 6.
38. Miller, G.A. *Language and Speech*; Freeman: San Francisco, CA, USA, 1981.
39. Doyle, L.R.; McCowan, B.; Hanser, S.F.; Chyba, C.; Bucci, T.; Blue, J.E. Applicability of information theory to the quantification of responses to anthropogenic noise by Southeast Alaskan humpback whales. *Entropy* **2008**, *10*, 33–46.
40. Sethares, W. *Tuning, Timbre, Spectrum, Scale*; Springer: London, UK, 1998.

© 2009 by the author; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license <http://creativecommons.org/licenses/by/3.0/>.