Review

Visual and acoustic communication in non-human animals: a comparison

GIL G ROSENTHAL and MICHAEL J RYAN*

Section of Integrative Biology C0930, University of Texas, Austin, Texas 78712, USA

Corresponding author (Fax, 512-471-9651; Email, mryan@mail.utexas.edu).

The visual and auditory systems are two major sensory modalities employed in communication. Although communication in these two sensory modalities can serve analogous functions and evolve in response to similar selection forces, the two systems also operate under different constraints imposed by the environment and the degree to which these sensory modalities are recruited for non-communication functions. Also, the research traditions in each tend to differ, with studies of mechanisms of acoustic communication tending to take a more reductionist tack often concentrating on single signal parameters, and studies of visual communication tending to be more concerned with multivariate signal arrays in natural environments and higher level processing of such signals. Each research tradition would benefit by being more expansive in its approach.

1. Introduction

Perhaps because of their primacy in the perceptual world of humans, vision and hearing are the two modalities most studied in animal communication. With a few notable exceptions (e.g. Julesz and Hirsh 1972; Dantzker *et al* 1999; Partan and Marler 1999) research in each area has followed on a fairly separate course. Differences in the two modalities, and in the approaches that workers have taken to each, have led to somewhat divergent conceptual models of how communication operates. In this review, we briefly compare the physical structure of visual and acoustic signals and how they are processed by receivers. We contrast the approaches taken to studying the two communication systems, and conclude by suggesting how scholars of each might benefit from considering the other.

2. Physical characteristics of signals

Identification, the first step in studying a signal, is usually more straightforward with acoustic systems. We take a signal to be any trait that has evolved specifically as a consequence of its effect on receiver behaviour (Bradbury and Vehrencamp 1998). Acoustic signals are most often produced in specific contexts, by organs whose primary function is the production of sound. In contrast, many morphological and chromatic visual traits are expressed continuously, independent of the signaler's motivational state or the environmental context. With these traits, it is often difficult to exclude functions other than communication. Long tails in male birds, for example, can serve as visual cues in female mate choice (Andersson 1982) but may also play a role in flight aerodynamics (Balmford *et al* 1993). Much of the work on visual traits, therefore, involves simply determining whether or not they are involved in communication.

We do not address the obvious differences in the means whereby acoustic and visual signals are produced. The great diversity of mechanisms within one modality alone makes it difficult to make general comparisons between the two; readers should refer to Bradbury and Vehrencamp (1998) for a survey of signal production mechanisms. The one generalization that can be made is that for sound, the energy reaching the receiver is nearly always generated by the signaler. Animals must produce and modify acoustic vibrations and then couple them to the surrounding medium. Highly specialized structures have evolved to accomplish these tasks.

With some exceptions (bioluminescent insects and marine organisms), the energy in visual communication is produced by the sun, reaching the receiver after reflecting off the signaler. The visual signal involves modifying available light to produce particular spectral, spatial, and temporal patterns. In some cases, such as the inflating dewlaps of many birds and lizards, visual signals are produced by specialized morphological adaptations, pigment cells, and musculature. One of the difficulties involved in identifying visual signals, however, is that visually conspicuous phenotypes can be generated for reasons that have nothing to do with communication; consider the brilliant colours of autumn leaves or dead crustaceans. [This difficulty is illustrated by the fact that the late, great evolutionary biologist, W D Hamilton, is reported to have suggested that brightly coloured leaves evolved as an aposematic signal to insects (*New York Times*, 10 March 2000 p A18).]

Both the spectral and temporal properties of an acoustic signal can be distorted or degraded by the environment intervening between the signaler and the receiver. Propagation media (air, water, substrate) absorb sound energy, with the proportion absorbed usually increasing with frequency. Objects with an acoustic impedance different from the medium, such as vegetation in terrestrial forests, scatter energy; again, higher frequencies are lost at higher rates. Reflection of acoustic energy off boundaries (e.g. the ground or the surface of a body of water in terrestrial systems) can produce attenuation of certain frequencies and augmentation of others, as can refraction in "sound channels" produced by gradients in temperature and pressure. Ambient noise can also produce degradation of signals; wind produces low-frequency noise, while other animals may produce high-frequency noise. Abiotic noise in oceans peaks at slightly higher frequencies (500 Hz) than does wind-produced noise on land, produced by weather at the surface and by water movements at depth (Rogers and Cox 1988).

In addition to frequency cues, temporal information in acoustic signals may be distorted due to reverberation, a problem especially acute in terrestrial forests and in aquatic environments, and by the introduction of amplitude modulations due to changes in the medium over the course of signal transmission (reviewed in Bradbury and Vehrencamp 1997; Ryan and Kime 2000).

Distortion and degradation of visual signals is the result of rather different processes. There are several important contrasts between the two modalities. The first, and perhaps most obvious, is that light requires line-of-sight transmission. The wavelengths of light used in visual communication are far smaller than many objects, so natural features lying between a signaler and a receiver will block transmission of visual signals. Light travels in straight lines, and cannot bend around objects in its path.

A second is that light travels far faster than sound, so differences in the speed of transmission between mediums and across the spectrum are largely irrelevant to visual perception. Doppler shifts, which can be used in some acoustic functions, such as echolocation, to gauge changes in position of a sound source, do not play an important role in visual signaling. Temporal information can thus be transmitted with greater fidelity in visual communication systems, although the temporal resolution of visual systems is lower than that of acoustic systems (Julesz and Hirsh 1972). Third, the efficacy of a visual signal depends on contrast with the background. The detectability of an acoustic signal is almost always optimized if it is emitted in an otherwise silent environment, in essence if there is no acoustic background. With visual signals, detectability depends on the contrast between the signaler and its background, and among internal components of the visual signal. Backgrounds that approximate the spatial, temporal, and spectral patterns of the signaler will render it effectively invisible (Endler 1980, 1984, 1990).

Fourth, the visual signal depends on the spectral and spatiotemporal quality of light which it reflects. The greatest effect of the environment on the perception of visual signals stems from variation in the way that sunlight is filtered before reflecting off the signaler. The spectral distribution of incident light depends on time of day and season, on meteorological conditions, and on the type and amount of vegetative cover. In aquatic systems, the incident light distribution changes as a function of depth; longer wavelengths attenuate fastest, so that shorter wavelengths (blues) predominate in deeper waters (Lythgoe 1988; Kirk 1994). Vegetation, such as a forest canopy, will also selectively filter certain wavelengths (Endler 1993a, b). The environment can also secondarily reflect light onto a signaler. Shallow, sandy sea bottoms, for example, reflect middle wavelengths onto the underside of fishes swimming midwater. The spatiotemporal properties of incident light can vary – consider the "dappling" of light in a forest understory. Ripples or waves on the water surface can also cause a focusing effect, causing illuminated objects underwater to flicker.

Between signaler and receiver, a visual signal can be distorted due to differential filtering by the medium of light at different wavelengths, a problem particularly in aquatic systems. Suspended particles in water, or fog on land, can form veiling light, which can reduce the contrast between an object and its background. In these cases, signals are distorted by scattering energy, as with frequency-specific attenuation in acoustic systems. Refraction of light, and reflection of light signals off sources between signaler and receiver, play less of a role in visual communication than does refraction and reflection in acoustic communication, except in special cases involving the air-water interface (Bradbury and Vehrencamp 1997).

The main generalization we can make about environmental effects on visual versus acoustic signals is that for acoustic signals, a far higher proportion of these effects come into play between the signaler and the receiver. The form of an acoustic signal that reaches a receiver depends on how it is produced by the signaler (which may depend on motivational state and on exogenous variables like temperature) and on how it is subsequently modified by the environment. With visual signals, there is an extra step at the beginning of this process, which is that the environment (and signaler behaviour in selecting light conditions) determines available light characteristics. The last step, modification of the signal between signaler and receiver, is comparatively less important in visual communication, particularly in terrestrial systems.

3. Sensory biology of receivers

We restrict our discussion to the lensed, image-forming eyes and pressure-detecting ears typical of vertebrates. Most far-field sound reception is accomplished through pressure detectors or pressure-gradient detectors, in which the displacement of a membrane by pressure triggers a sensory receptor. In vertebrates that hear in air, a tympanic membrane in the outer ear is directly coupled to the surrounding medium. Pressure variation of sound causes movement of this membrane; the mechanical movement is transduced and amplified by a series of small bones in the middle ear to movement of fluid in the inner ear canal. The fluid motion causes the movement of certain inner ear membranes and displacement of sensory hairs cells that rest upon them. These hair cells are often organized tonotopically – different parts of the membrane are stimulated by different frequency intervals. Displacement of the hair cells results in neural discharges and, by some definitions, hearing (Fay and Popper 1994, 1999).

Aquatic organisms are acoustically transparent; the auditory system of most fishes uses a gas-filled swimbladder and/or mineralized otoliths to generate differences in acoustic impedance. Like in terrestrial organisms, the receptor organ which transduces acoustic energy is tonotopic (Fay and Popper 1994, 1999).

In vision, a lens projects a two-dimensional image onto the retina. An array of photoreceptors then converts the light information into neural signals. Photoreceptors, or cones, contain visual pigments which undergo a conformational change when excited by a photon. The retinas of animals with colour vision contain multiple (2–4) classes of cones, each sensitive to a specific region of the visible spectrum. The sensation of colour results from a colour-opponent system, based on excitatory and inhibitory inputs from each type of cone photoreceptor. The colour-

opponency system is capable of discriminations as fine as 2 nm (Jacobs 1981). Nevertheless, for a trichromatic system like that of humans, every colour that we see can be represented by a linear combination of three pure wavelengths, a fact exploited by photography and video.

The two-dimensional retina preserves spatial relationships, except for depth, in the projected image. Edges are detected in the retina itself, using a system of excitatory and inhibitory inputs from photoreceptors similar to that used for colour detection. Spatial-frequency processing uses Fourier analysis much as frequency processing does in hearing. Cells at higher levels of processing, such as in the visual cortex of primates, are tuned for particular spatial frequencies. Other units respond to specific orientations or temporal frequencies of stimuli (Sekuler and Blake 1994).

The visual system uses a variety of means to perceive depth. Binocular disparity – the discrepancy between projected areas on two different retinas – can provide cues of distance to an object (Pettigrew 1991). Higher-order monocular cues include accommodation cues, parallax, occlusion, and texture gradients as indicators of depth (Bradbury and Vehrencamp 1998).

Acoustic and visual systems differ in several important ways. Visual systems offer much higher spatial resolution, producing a detailed map of a receiver's environment (Julesz and Hirsh 1972). Vision provides direct information on the two-dimensional spatial relationships between objects, and to a lesser extent on depth cues. Hearing relies on binaural cues to determine the location and distance of a sound source, and spatial resolution is poor. Temporal relationships, in contrast, are more finely detected by acoustic systems (Julesz and Hirsh 1972), although they are more likely to be distorted during acoustic transmission. The inner ear often contains numerous receptors narrowly tuned to specific frequency bands, while colour vision relies on analysis of the response of a few broadly tuned receptors. Julesz and Hirsh (1972), in a review concentrating on human sensation, characterized the visual system as providing information on an animal's general environment, while the acoustic sense was dedicated to evaluating communication signals. While this generalization holds true for animals like frogs, humans, and most birds, there are many animals, including many reptiles and fishes, for which vision may play a dominant role in communication.

4. Modality and the evolution of communication systems

Differences in the physical characteristics of signals and how they are perceived are likely to play a major role in

the evolution of both signalers and receivers. At one extreme, a sensory modality has evolved primarily in the context of communication, as appears to be the case for hearing in frogs. This has the effect of relaxing natural selection on the modality and allowing for tuning towards properties characteristic of conspecifics. Among frogs, there is considerable variation in the complexity of one of the two inner ear organs sensitive to air-borne sound, the amphibian papilla (AP). It has been suggested that increased complexity of the AP yields a greater range of sensitivity, which promotes greater call diversity and the potential for divergence of species recognition systems that serve as reproductive isolating mechanisms. Among the major clades of anurans, there is a strong correlation between complexity of the AP and number of species, supporting the notion that complexity of the receiver in the species recognition system promotes speciation (Ryan 1986).

At the other extreme, a sensory modality may have evolved as the result of an array of selective forces in addition to communication. Most visual systems appear to reflect a compromise necessary to perform a variety of functions such as prey detection, predator avoidance, and conspecific recognition (Endler 1984). Interestingly, Land and Fernald (1992) suggested that the evolution of image-forming eyes is correlated with greater rates of speciation due to the general utility of vision in the animal's ecology rather than any role in communication of species recognition information being responsible for the greater speciation rates (see also Lythgoe 1979; Lythgoe *et al* 1994). de Queiroz (1999), however, conducted a phylogenetic comparision of similar animals with and without eyes and did not find support for this hypothesis.

Visual systems are in general broadly tuned, with considerable overlap among species in visual sensitivity. In acoustic communication, a signal can be highly conspicuous to one species and yet undetectable to another, simply because animals vary widely in their range of sensitivity to acoustic frequencies (Heffner and Heffner 1980). Increasing the amplitude of a call can make it substantially more detectable to a conspecific. If the call's frequency envelope lies outside the frequency sensitivity range of a predator, this increase carries no cost: the predator remains deaf to the call. Vision, however, is distinctive in that signal detection relies on the relative contrast between a figure and its background, rather than the absolute brightness of the figure. Since the perception of visual signals depends on detecting background contrast, an object outside an observer's range of spectral sensitivity will appear black, which may be highly detectable against many natural backgrounds. Increasing the apparent size of the object will result in a larger, and therefore more detectable, black object. There are few ways in which a visual signal can, like a "private" acoustic signal,

be rendered completely undetectable to potential enemies while remaining conspicuous to intended receivers. One strategy is to minimize the internal contrast of ornaments with background body colouration (Endler 1991).

Visual systems tend to be adapted to environmental light characteristics (Lythgoe 1979), suggesting that animals in similar habitats will often share similar visual perception. Due to the physical properties of light, moreover, visual perception is confined to a narrow range of wavelengths, between 300–700 nm. Most animals, even those adapted to low-light, restricted-bandwidth conditions (e.g. Yokoyama *et al* 1999) are sensitive over at least half this range; with acoustic signals, sensitivity ranges vary overwhelmingly within mammals alone: consider the "subsonic" world of elephants, the "sonic" world of humans, and the "ultrasonic" world of bats (Heffner and Heffner 1980; Fay and Popper 1994).

The evolution of both acoustic and visual signals is constrained by environmental effects on signaling. In acoustic systems, degradation and distortion between the sound source and the receiver can exert selection on signal design (Bradbury and Vehrencamp 1998; Ryan and Kime 2000). In visual systems, available light (Lythgoe 1979; Endler 1993b; Partridge and Cummings 1999) and the spatiotemporal characteristics of background (Fleishman 1992) can structure the evolution of visual signals.

5. Studying communication

While acoustic and visual communication each involve multiple physical dimensions, research has concentrated on only a subset of these. The overwhelming majority of research on animal acoustic communication has concentrated on the frequency and temporal properties of signals, although a number of neuroethological studies have addressed localization of a sound source (Konishi 1993) and some behavioural studies have addressed the spatial properties of emitted sounds (Dantzker *et al* 1999). Considerations of cognitive aspects of audition (Bregman 1990; Mellinger and Mont-Reynard 1995) emphasize the fact that auditory environments consist of complex scenes, which auditory systems must analyze much as visual systems do visual scenes. Auditory studies have generally not addressed the role of these higher-level properties in animal acoustic communication.

Research on visual communication falls into two broad classes: work on higher-level stimulus properties like form, pattern, and motion, and work on colour. A handful of studies (Fleishman 1992) have explicitly looked at spatial (Endler 1984) and temporal properties. These studies and those on colour have often followed the same methodological rigour as many acoustic studies, in which the

physical properties of a signal are explicitly quantified. Most studies examining higher-level properties, however, do not quantify the colour, spatial, or temporal properties of stimuli, backgrounds, or ambient light conditions, whether characterizing signals themselves or evaluating receiver response.

These limitations are in part due to the difficulties involved in quantifying the parameters in question. Obtaining an accurate recording of an acoustic signal emitted by a point source, a point estimate of reflectance, or a video sequence of a two-dimensional scene is straightforward. Obtaining spatial information via a three-dimensional array of synchronized microphones, a two-dimensional distribution of reflectances, or a three-dimensional sequence of a visual scene is not.

Nevertheless, there are clear benefits to considering the approach taken in studies of one modality when studying the other. Acoustic studies could address the acoustic scene – particularly the features of signals and receivers that determine selective attention and parsing of signals from background in a complex natural acoustic environment. Visual studies could strive for more rigorous characterization of the visual scene for complex stimuli, including quantification of natural backgrounds and ambient light distributions.

The purpose of this essay was to note differences between communication in visual and auditory modalities, contrast the differences in how these systems are studied, and to encourage a broader approach within each. While the physical nature of acoustic and visual stimuli leads to divergent strategies for detection and processing, there are many similarities in the type of information conveyed and in the way this information is analysed. An awareness of how animals operate in different modalities could yield valuable insight into the structure and evolution of signal-receiver systems.

Acknowledgements

References

Andersson M 1982 Female choice selects for extreme tail length in a widowbird; *Nature (London)* **299** 818–820

Balmford A, Thomas A L R and Jones I L 1993 Aerodynamics and the evolution of long tails in birds; *Nature (London)* **361** 628–631

Bregman A S 1990 *Auditory scene analysis, the perceptual organization of sound* (Cambridge, Massachusetts: The MIT Press)

Bradbury J W and S L Vehrencamp 1998 *Principles of animal communication* (Sunderland, Massachusetts: Sinauer)

Dantzker M S, Deane G B and Bradbury J W 1999 Directional acoustic radiation in the strut display of male sage grouse *Centrocercus urophasianus*; *J. Exp. Biol.* **202** 2893–2909

de Queiroz A 1999 Do image-forming eyes promote evolutionary diversification?; *Evolution* **53** 1654–1664

Endler J A 1980 Natural selection on color patterns in *Poecilia reticulata*; *Evolution* **34** 76–91

Endler J A 1984 Progressive background matching in moths, and a quantiative measure of crypsis; *Biol. J. Linn. Soc.* **22** 187–231

Endler J A 1993a The color of light in forests and its implications; *Ecol. Monogr.* **63** 1–27

Endler J A 1993b Some general comments on the evolution and design of animal communication systems; *Philos. Trans. R. Soc. London B* **340** 215–225

Endler J 1990 On the measurement and classification of colour in studies of animal colour patterns; *Biol. J. Linn. Soc.* **41** 315–352

Endler J A 1991 Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions; *Vis. Res.* **31** 587–608

Fay R R and Popper A N (eds) 1994 *Springer handbook of auditory research: Comparative hearing: Mammals* (Berlin: Springer Verlag)

Fay R R and Popper A N (eds) 1999 *Springer handbook of auditory research: Comparative hearing: Fish and amphibians* (Berlin: Springer Verlag)

Fleishman L J 1992 The influence of the sensory system and the environment on motion patterns in the visual displays of anoline lizards and other vertebrates; *Am. Nat.* **139** S36–S61

Heffner R and Heffner H 1980 Hearing in the elephant (*Elephas maximus*); *Science* **208** 518–520

Jacobs G H 1981 *Comparative color vision* (New York: Academic Press)

Julesz B and Hirsh I J 1972 Visual and auditory perception – an essay of comparison in human communication; in *Human communication: A unified view* (eds) E E David and P Denes (New York: McGraw-Hill) pp 283–335

Kirk J T O 1994 *Light and photosynthesis in aquatic ecosystems* 2nd edition (Cambridge: Cambridge University Press)

Konishi M 1993 Neuroethology of sound localization in the owl; *J. Comp. Physiol. A* **173** 3–7

Land M F and Fernald R D 1992 The evolution of eyes; *Annu. Rev. Neurosci.* **15** 1–29

Lythgoe J N 1979 *The ecology of vision* (Oxford: Clarendon Press)

Lythgoe J N 1988 Light and vision in the aquatic environment; in *Sensory biology of aquatic animals* (eds) J Atema, R R Fay, A N Popper and W N Tavolga (New York: Springer-Verlag) pp 57–82

Lythgoe J N, Muntz W R A, Partridge J C, Williams J and Shand D M 1994 The ecology of the visual pigments of snappers (Lutjanidae) on the Great Barrier Reef; *J. Comp. Physiol.* **A174** 461–467

Mellinger D K and Mont-Reynard B M 1995 Scene analysis; in *Springer handbook of auditory research: Auditory computation* (eds) H L Hawkins, T A McMullen, A N Popper and R R Fay (Berlin: Springer Verlag) pp 271–331

Partan S and Marler P 1999 Communication goes multimodal; *Science* **283** 1272–1273

Partridge J C and Cummings M E 1999 Adaptations of visual pigments to the aquatic environment; in *Adaptive mechanisms in the ecology of vision* (eds) S N Archer, M B A Djamgoz, E R Loew, J C Partridge and S Vallerga (Dordrecht: Kluwer Academic Pub.) pp 251–283

Pettigrew J D 1991 Evolution of binocular vision; in *Evolution of the eye and visual system* (eds) J R Cronly-Dillon and R L Gregory (Boca Raton: CRC Press) pp 271–283

Rogers P H and Cox M 1988 Underwater sound as a biological stimulus; in *Sensory biology of aquatic animals* (eds) J Atema, R R Fay, A N Popper and W N Tavolga (New York: Springer-Verlag) pp 131–149

Ryan M J 1986 Neuroanatomy influences speciation rates among anurans; *Proc. Natl. Acad. Sci. USA* **83** 1379–1382

Ryan, M J and Kime N M 2000 Selection on long distance acoustic signals; in *Springer handbook of auditory research; Acoustic communication* (eds) A Simmons, R R Fay, and A N Popper (Berlin: Springer Verlag) (in press)

Sekuler R and R Blake 1994 *Perception* (New York: McGraw-Hill)

Yokoyama S, Zhang H, Radlwimmer F B and Blow N S 1999 Adaptive evolution of color vision in the Comoran coelacanth (*Latimeria chalumnae*); *Proc. Natl. Acad. Sci. USA* **96** 6279–6284

Corresponding editor: RENEE M BORGES