Acoustics of Bird Song—Some Unresolved Problems

The present state of our understanding of the acoustics of bird song is reviewed from the viewpoint of a physical scientist. It is concluded that the standard model for the production of voiced song, in which the vibrating syringeal membranes act as a pressure-controlled flow valve, is adequately established, in that it is possible to explain most of the features observed in such song with a good measure of quantitative success. The mechanism for production of whistled song is not yet established, but three alternative proposals are put forward in sufficient detail that they can now be subject to anatomical verification and, if not ruled out on such grounds, used as the basis of a quantitative model. The most attractive of these proposals involves excitation of the syringeal membrane by a wave-like disturbance on a jet flowing tangential to it. The resonant properties of the interclavicular air sac may also be involved.

INTRODUCTION

In keeping with the aims of this journal, the paper that follows is a personal view of the current state of our knowledge of the acoustics of bird song, and of the most pressing problems that remain to be solved. The approach is quite frankly that of a physical scientist, rather than of a biologist, so that the emphasis may be unfamiliar. I make no apology for this, and indeed I have taken the view that it is better to present a perspective that is complementary to that of my biological colleagues rather than to try for a carefully balanced exposition. In doing this I do not wish to imply

Comments Theoretical Biology 1989, Vol. 1, No. 4, pp. 237–251 Reprints available directly from the publisher Photocopying permitted by license only © 1989 Gordon and Breach, Science Publishers S.A. Printed in Great Britain that the physical approach is in some way the better one, but merely that it is different. By examining a complex system from several different angles, with different backgrounds of familiar knowledge, we can perhaps most easily make progress.

Bird song has been of interest to scientists for a very long time, and published studies go back more than a century. Excellent surveys are given in the classic volume by Greenewalt¹ and in recent reviews by Brackenbury² and by Gaunt and Gaunt.³ These publications contain copious references to the primary literature.

Although the variety of bird vocalization is immense—understandably much greater than the variety of human vocal utterance, which derives from only a single species—it is useful to make some human analogies. Birds certainly produce transient sounds rather similar in nature to the consonants of human speech, and there does not seem to be a great deal of mystery about their acoustic origin, though the details may be complex. Setting these transients aside, birds also produce sustained sounds which are somewhat analogous to human voiced vowels and whistles, and it is these song features which attract most interest and which will concern us here.

As documented by Greenewalt¹ by means of Sonagrams and detailed measurements, these sustained sounds can be conveniently divided into three classes: "voiced" sounds, consisting of a wide spectrum of harmonically related components distributed under an envelope with one or more formant bands, narrow-band sounds with several components equally spaced in frequency but not necessarily harmonically related, and "whistled" sounds consisting of a single pure-tone component with the level of its overtones being below -20 dB or even -30 dB relative to the fundamental.⁴ In some species, the pitch, or general frequency level, of the song varies very little, while in other species the fundamental frequency varies by as much as a factor 4, giving a pitch range of 2 octaves or more. It is also established that many birds have two independent vocal sources which can operate simultaneously.

It is my view that the fundamental mechanism by which "voiced" song is produced is now moderately well understood, both qualitatively and quantitatively. In contrast, however, there is not even qualitative agreement on the mechanism of "whistled" song, and certainly no quantitative description. The narrow-band song is,

perhaps, usually a variant of whistled song, so that its origin, too, is obscure, though it can be understood in terms of amplitude or frequency modulation of a pure-tone whistled song.

I propose here to review briefly the current understanding of the acoustics of voiced song and to use this as a starting point from which to suggest two possible specific mechanisms for whistled song. In doing this, I recognize that it is of little value to simply offer hand-waving speculations, so that my proposals are couched in quantitative terms, and suggestions are made for their investigation.

THE ACOUSTICS OF VOICED SONG

The general anatomy of the vocal system of passerine birds has been described in detail by Warner,⁵ though of course much information is also available in the earlier literature. Greenewalt¹ reviews publications back to the time of Hérissant in 1753, which show that the form and function of the avian syrinx were basically understood as long ago as that, at least by analogy with woodwind reed instruments.

For our present discussion we do not need to concern ourselves with minor anatomical variations, and we can be content with a very generalized model for the vocal system, as is shown in Fig. 1. The active part of the system consists of sidewall membranes in the two bronchi below the point at which these merge to form the trachea. These membranes, relaxed during normal respiration, can be made to protrude into the bronchial airways by the action of air pressure in the surrounding interclavicular air sac, modified by membrane tension provided by muscular action. The opposing cartilage may also be moved by muscular action in some species to further constrict the airway. As remarked above, this vocal system has two apparently separate sources, in distinction to the human system in which the vibrating membranes form a pair of vocal folds in the trachea above its junction with the bronchi.

There is little argument but that bird song is associated with flow of air past these syringeal membranes, and that in most cases the membranes vibrate during vocalization. Until recently, quantitative studies have been lacking at the mechanistic level. Data

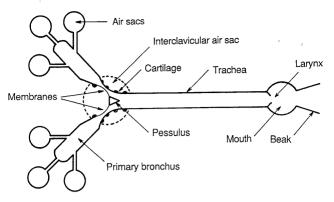


FIGURE 1 Schematic view of the vocal system of a typical bird. Details of the implied anatomy are not important. (After Fletcher, with minor modifications).

are available, however, on typical air flow rates, internal pressures, and acoustic power outputs for a variety of birds. Brackenbury⁶ has measured typical acoustic powers for 18 different birds, ranging in body mass from 6 g to 3.5 kg. Not surprisingly, acoustic power varies widely, from about 0.2 mW for the smallest birds, through common values in the range 1 to 10 mW for medium sized birds, to a peak as high as 200 mW for the short cries of the common rooster *Gallus domesticus*.

Measurements of physiological variables such as pressure and air flow during song are less readily available, but particular cases have been reported by Brackenbury⁷ and by Gaunt and Gaunt.³ Again the variation between species is large, but typically the air sac pressure during song (which should be essentially the same as the pressure in the bronchi below the syrinx) is about 100–200 Pa in small birds, and as high as 1,000 Pa in *Gallus domesticus*. The pressure in the trachea is typically lower than that in the sac by a factor 3 to 10, so that there is a large pressure drop across the syrinx. Air flow rate depends greatly upon the size of the bird, and ranges from about 10 ml/s to as much as 500 ml/s, this extreme value again referring to *Gallus domesticus*, though not at its peak loudness. Some birds apparently inhale during part or even all of their song.

From these figures we can calculate the overall efficiency with which pneumatic power is converted to acoustic power, at least

for Gallus domesticus for which sufficient data are available. This efficiency is about 3%, which is comparable with, though rather greater than, that of sound production in musical wind instruments.⁸ Much of the power loss occurs in flow through the syrinx constriction, and the remainder in viscous and thermal losses to the walls of the trachea.

Until recently there does not appear to have been any attempt to combine the available anatomical and physiological information into a quantitative model for the mechanism for voiced song. Indeed the closest approach is probably that in Greenewalt's book, which does treat some parts of the acoustic system in a semiquantitative way. Most other discussions have been of a qualitative nature and have assumed that the Bernoulli force is the major driving mechanism, completely ignoring both the large pressure drop across the syringeal constriction and the importance of phase relationships in determining whether energy is supplied by, or lost to, the flow. Similar comments apply to the role of tracheal resonances, which cannot be settled by analogy with either musical instruments or with the human voice.

It turns out to be quite straightforward to put together all the available information into a quantitative model⁹ for the functioning of the vocal system—we might term this the "standard model," since it simply quantifies the proposals that have been available in the literature since the time of Greenewalt. This model gives a very good account of the mechanism of voiced song, matching the general properties of the observed spectrum for a typical bird, and giving a power output in good agreement with observation for typical values of the anatomical and physiological variables. It clarifies the origin of the exact harmonic spectrum in terms of nonlinear coupling of the inharmonic membrane modes, and sets to rest the worries that have been expressed by some writers¹⁰ on this score. The model also accounts quantitatively for the role of both pressure-drop and Bernoulli terms in driving the syringeal membranes, without introducing additional assumptions.

The various formants imposed by the tracheal resonances are also correctly modeled, though it should be remarked that the presence or absence of "open-pipe" resonances depends upon the acoustic impedance of the respiratory system below the syrinx, a point on which there does not appear to be any present information. The possible role of the larynx and mouth in modifying the frequencies of the tracheal resonances, and thus of the formant bands, is mentioned, but has yet to be explored in detail. While the lowest natural frequency of the syringeal membranes, backed by air-sac pressure, appears as the determinant of song pitch, the model allows interaction of tracheal resonances with this membrane resonance if the frequencies are comparable, as in birds with high-pitched voices relative to their size.

It is important to emphasize that, in this model, the syringeal membranes do not act as direct radiators of sound—they simply serve as the components of a pressure-controlled valve which modulates the flow of air through the syringeal contraction. The acoustic power is carried by this modulated air flow. We return to this point in a later section.

Although the model was developed for a simplified one-voice situation, it is straightforward enough to extend it to treat the real case in which the bird has vocal sources in each of its bronchi. The equations of the model are then supplemented by a second similar set, and the only coupling terms between them are provided by the acoustic pressure at the base of the trachea. This coupling is weak, compared with the modal coupling in an individual membrane caused by flow nonlinearity, so that the two voices are able to act, to a large degree, independently. Coupling may occur, however, if their individual frequencies are close together.

The model can be applied to birds of any species by appropriate choice of the relevant anatomical and physiological parameters, and appears capable of explaining most features of voiced song, though such exploration has not yet been carried out. We should not, of course, expect detailed agreement because of the simplified nature of the model itself. The simplified nature of the model should be emphasized, indeed, not as a criticism, but rather as a virtue, in that it reduces the number of unknown parameters available for manipulation. This should be contrasted with the current state of development of models of the human larynx, 11 which may use no less than 16 coupled masses for each vocal cord, an 18 section horn for the pharynx, and a 12 section nasal tract!

Brief exploration of the capabilities of the model has shown that it is not able to produce sound with very low harmonic content. Greenewalt¹ had suggested that such song might be produced if

the vibration amplitude of the membranes were to be small enough that they failed to touch the opposite side of the airway. In fact it is very difficult to produce a stable oscillation of this type—the membrane amplitude either grows until contact occurs or else dies away—and in any case the harmonic content is quite large even for moderately small vibration amplitudes because of the considerable nonlinearity of the flow. We must therefore look elsewhere for a model for "whistled" song.

A POSSIBLE AIR-SAC RESONANCE

The interclavicular air sac is important in vocalization, because it is the pressure in this sac which causes the syringeal membranes to protrude into the airway during song. This pressure, along with the mechanical properties of the membranes themselves, also determines the membrane resonance frequencies, as discussed in the standard model.⁹ It is possible, however, that the air-sac cavities play a more active role than this, by contributing additional resonances to the system. Let us examine this first in relation to voiced song.

The air sacs are connected directly to the respiratory system by apertures and tubes, in most species, and this ensures that the steady sac pressure is normally the same as that in the bronchi. Because the connecting apertures are small, however, it is possible for the acoustic pressure in the sacs to be different from that in the rest of the respiratory system, provided that the frequency involved is above the inverse relaxation time for acoustic flow between the sac and the lungs. In particular, this allows the possibility of resonances for the system consisting of the interclavicular cavity loaded by the syringeal membranes. For a bird of medium size, the sac volume might be 10 cm³, the membrane area 0.1 cm², and the membrane thickness about 10 microns. These dimensions give a cavity resonance frequency of about 500 Hz, provided the membrane frequency is lower than this so that it acts as an inertive load. There is, of course, a good deal of uncertainty about this estimate, and the bird might change the frequency over a large range by changing the sac inflation, the relaxation time for which is probably at least 0.1 s.

The possible existence of such a sac resonance creates a further formant frequency which might be under voluntary control and which could assist in providing this formant in "talking" birds. 12 This speculation could be tested by experimental examination.

POSSIBLE MECHANISMS FOR "WHISTLED" SONG

The dichotomy implied by using separate names for voiced and whistled song begs the question whether they are, in fact, acoustically distinct, or simply extreme versions of song produced by a common mechanism of the type discussed above. The extremely low harmonic content of whistled song, despite the fact that its radiated power is typically comparable with that of voiced song, argues for a separate mechanism, given the inherently large nonlinearity of a syringeal membrane-valve source. The fact that the model discussed above, while admittedly primitive, either produces richly harmonic sound or no sound at all, supports this contention. It is therefore useful to examine other possible mechanisms that might be able to account for the acoustic features of whistled song. The first of these is a variant of the mechanism for voiced song, while the others are entirely different. We can also distinguish between syringeal mechanisms, in which the acoustic output is necessarily influenced by the acoustic properties of the trachea, and oral mechanisms, in which the sound is produced in the mouth and is therefore free from tracheal influence.

A Membrane-related Mechanism

While we have seen that vibrating syringeal membranes, acting as the components of a pressure-controlled valve in a narrow constriction of the airway, constitute a very nonlinear acoustic source, and are thus incapable of producing pure-tone song, it is worthwhile to investigate other possible mechanisms in which the membranes act as flow-excited direct radiators. A proposal of this kind—but without explicitly recognizing its difference from the mechanism of the standard model—was put forward by Brackenbury. ¹³ His membrane was supposed to be located in a constriction in the airway, to be excited solely by Bernoulli forces, and to have a displacement amplitude small relative to the airway dimension,

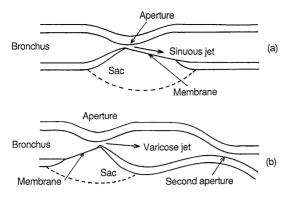


FIGURE 2 Schematic view of two possible avian vocal systems which seem capable of producing whistled song: (a) a system with a constriction upstream from the syringeal membrane, producing a tangential jet whose unstable sinous instabilities couple to displacements of the syringeal membrane, and perhaps to resonances of the underlying cavity; and (b) a system with an extra aperture, potentially capable of producing an aperture-tone whistle through varicose instability of the jet.

but these physical constraints—necessary in order to limit the volume flow to a reasonable value—could not be made consistent with a reasonable acoustic output.

I put forward here a rather different mechanism, based upon formation of an air jet at a constriction in the airway upstream from the membrane, which is itself located in a rather wide section of the airway, as shown in Fig. 2a. It is envisioned that these anatomical features might be fixed, in a bird singing only in the "whistle" mode, or produced by muscular adjustment of syringeal cartilage in those species able to produce both voiced and whistled song.

There are two important features associated with the presence of an air jet, rather than a simple flow, in the acoustic generator. The first is that the constriction forming the jet limits the flow, so that a reasonably large jet velocity, as determined by the Bernoulli flow equation, can be achieved without at the same time involving an unreasonably large volume flow from the lungs. The second is that an air jet emerging from a constriction is acoustically unstable, as analyzed long ago by Rayleigh, ¹⁴ and can serve as a source of excitation for a resonant system. An additional feature of this model is the fact that, being located in a wide section of windway,

the membrane can reach a large vibration amplitude without impeding the airflow and generating harmonics through nonlinear effects.

If we assume a typical bronchial pressure of about 200 Pa, then the airspeed in the jet will be about 20 m/s, and the wave speed on the jet^{15,16} will be about half this value, or 10 m/s. If the jet flows tangentially across the membrane then the membrane will serve simply to define one face of the jet, and will be influenced by the instability of the jet itself. Any small displacement of the membrane will displace the jet near the aperture, and this displacement will be amplified by aerodynamic effects as the displacement wave travels along the jet. The situation is somewhat different from the normal arrangement in flutes and organ pipes, where there is an opening, rather than a thin membrane, across which the jet travels, but an analysis similar to that for the organpipe case¹⁶ shows that self-sustained oscillation will occur when the length of jet travel across the membrane is about one half wavelength of the disturbance on the jet, though there is a good deal of latitude in this dimension, particularly on the short side. There is no threshold jet speed, as in Brackenbury's case, apart from this requirement. From the typical figures given above, this requires a jet length of about 5 mm if the sound frequency is about 1000 Hz, which is anatomically reasonable. If the membrane has a reasonably sharp resonance at this frequency, whether of itself or due to the influence of modes in the backing cavity as discussed above, and if its amplitude of motion is small enough to not restrict airflow, then its motion should be nearly sinusoidal. The oscillation frequency can be controlled by changing the tension in the membrane, the degree of inflation of the underlying sac, or both. There would probably be an increase of air pressure associated with high frequency notes.

In contradistinction to the sound source in the standard model, the membrane here acts as a simple vibrating radiator, loaded by the tracheal horn. The resonances of the trachea clearly have the potential to influence the radiated sound, but the magnitude of the effect depends on the impedances of the various parts of the system. An estimate ignoring tracheal resonances, similar to that made by Brackenbury, shows that a radiated power or order 1 mW can be achieved with a membrane amplitude of order

1 mm, which is again anatomically reasonable.

This mechanism is inherently appealing, since it represents only a small anatomical adjustment from the conditions associated with voiced song, and thus makes this vocal mode readily available to birds. The model depends, however, upon the necessary anatomical features or adjustments being present or possible. It is therefore possible to subject it to initial experimental assessment. Equally, once reasonable anatomical parameters have been established, it is not difficult to construct a quantitative model, along the lines of that discussed above for voiced song, and to calculate the acoustic output, both in power and harmonic content, for given physiological input.

A Syringeal Aerodynamic Mechanism

The only other plausible mechanism appears to be some sort of aerodynamic whistle, in which there is no essential vibration of solid surfaces and sound is produced by fluctuating air flow produced solely by aerodynamic forces. This possibility was not examined by Greenewalt, but has been raised since by Gaunt and Gaunt³ and by Casey and Gaunt.¹⁰ These authors are vague about the nature of the possible mechanism, except to postulate that it depends upon the generation of periodic vortices upon a jet of air emerging from a slit formed by the syringeal membranes. They note that whistle tones produced by supposedly similar mechanisms in other physical systems have the feature of being nearly pure tones, as is required to explain this characteristic of whistled bird song.

To bring this speculation closer to reality, it is necessary to be much more specific about the nature of the tone-generation mechanism. The study of aerodynamically generated sound, as with most things acoustic, goes back to the time of Rayleigh. A conveniently accessible description of whistle generators is given by Chanaud. Most qualitative discussions focus rather too much on vortices, which are the end-result of instability, rather than on jet instability and wave growth, which is the essential feature. A Chanaud points out, aerodynamic whistles are conveniently divided into 3 classes—those producing oscillation by simple jet instability or by aerodynamic feedback around an obstacle (aeolian tones), those relying on acoustic feedback from an edge or a second aperture (edge tones or aperture tones), and those relying upon

acoustic interaction between an air flow and an acoustic resonator. We have already discussed the last of these possibilities in the section above. There do not seem to be any other available resonators in the syrinx, and the tracheal resonances, being relatively fixed in frequency, do not seem to provide a useful alternative.

At ordinary air velocities, the sound level produced by the instability of a jet in the absence of obstacles and acoustic feedback is too small to produce a plausible model source. The only feasible syringeal source appears to be one consisting of a pair of apertures, separated by a distance equal to about half a wavelength of the jet disturbance, and thus typically about 5 mm, as shown in Fig. 2b. Such a system could be excited by varicose instability of the jet. Inspection of photographic sections through the syrinx⁵ does not show any obvious candidate structures, but the suggestion is perhaps worthy of more detailed study. In passing, we should note that, for this sort of aperture tone generator, the frequency of the tone is determined primarily by the jet velocity, and thus varies at the square root of blowing pressure, unless the distance between the two slits can be altered.

A Possible Oral Source

An oral source for whistled song has some attractiveness as a hypothesis, since this immediately avoids the complications of tracheal resonances and the difficulty they involve for the explanation of the production of smooth *glissandi* over a wide frequency range. There is also the convenient analogy with human whistling, which demonstrably produces a nearly pure-tone acoustic output which is readily controllable over a large frequency range by very small movements of the mouth parts.

In human whistling, the first of a pair of apertures is generally produced between the tongue and hard palate, and the second between the lips, though other configurations are possible. The larynx of the bird could provide a suitable jet-forming aperture, since its opening slit is variable, but there seems to be no candidate structure for a second aperture, given the rather inflexible nature of the mouth parts. It is possible that an edge-tone rather than an aperture-tone might be involved, but the mouth structures do not seem to provide sufficient flexibility for this to be the common sound-producing mechanism in birds.

Despite these objections, it is not sensible to rule out the possibility of an oral source without investigation, particularly since this could be carried out, at least in principle, by telephotographic observation coupled with sound recording.

CONCLUSIONS

The classic mechanism proposed for bird vocalization, involving the air-driven vibration of syrinx membranes which themselves control air flow, appears to be adequately supported by quantitative analysis as the mechanism for at least some types of bird song. The available quantitative models are, as yet, primitive and generic, and require extension and specialization if they are to give good descriptions of the voiced song of birds with a large range of anatomies and vocal patterns, but there does not seem to be any real barrier to progress in this direction. Such progress will be achieved, however, only by quantitative modeling of realistic anatomies, and by comparing the predictions of such models with physiological and acoustic measurements specifically designed to test them.

The origin of whistled song remains, in my view, an open question, but I have proposed above, in sufficient detail for them to be tested, two possible syringeal mechanisms and one oral mechanism which seem capable of producing controllable and nearly pure-tone sound. The first of these syringeal mechanisms is, I think, the most attractive. The point is, however, that it should now be possible to test these mechanisms in two entirely different ways.

In the first place, each of the syringeal mechanisms relies upon the possibility of a different particular conformation of the syrinx which should be achievable, since some birds sing in both ways, by a muscular shift of certain anatomical elements. For the jet-excited-membrane model, it must be possible to constrict the airway just upstream from the syringeal membranes in such a way that they are left in a wide part of the airway. For the aperture-tone model, it must be possible to produce a second aperture in the airway, downstream from the syringeal contraction. Both these suggestions can be investigated, at least in principle, by straight-

forward anatomical and physiological methods. The proposed oral mechanism can similarly be studied, in the first instance, by simple observation.

In the second place, now that the functioning of the three proposed models has been made explicit, it is possible to set up and investigate precise acoustic models for the vocal system when acting in this way. These models, if successful, should be able to produce the required level of acoustic power, and the required single-component acoustic spectrum, for reasonable values of all the relevant anatomical and physiological parameters, in the same way as did the model for voiced song. Indeed, it should be possible to produce a single model with an appropriately adjustable syrinx configuration, and perhaps mouth configuration as well, and to show that it can produce either voiced or whistled song.

This is not a simple program that I have outlined, but I think it is the only way in which to make progress. The models advanced are, I think, plausible, and none of them can be ruled out *a priori* without careful investigation. If one of the models, or even more than one, meets the tests I have proposed, then we will be much closer to an understanding of an important variety of bird song. If all fail, then we must seek a different model and subject it to the same sort of scrutiny.

N. H. FLETCHER

Division of Radiophysics, CSIRO Australia, and Research School of Physical Sciences Australian National University, Canberra ACT 2601, Australia

References

- 1. C. H. Greenewalt, *Bird Song: Acoustics and Physiology* (Smithsonian Institution Press, Washington, DC, 1968).
- J. H. Brackenbury, in Acoustic Communication in Birds, eds. D. E. Kroodsma, E. H. Miller and H. Ouellet (Academic Press, New York, 1982) Vol. 1, pp. 53-73.
- 3. A. S. Gaunt and S. L. L. Gaunt, in *Current Ornithology*, ed. R. F. Johnston (Plenum Press, New York, 1985) Vol. 2, pp. 213-245.
- 4. Ref. 1, p. 145.
- 5. R. W. Warner, J. Zool. Lond. 168, 381 (1972).
- 6. J. H. Brackenbury, J. Exp. Biol. 78, 163 (1979).

- J. H. Brackenbury, J. Exp. Biol. 72, 229 (1978).
 A. Bouhuys, J. Acoust. Soc. Am. 17, 453 (1965).
 N. H. Fletcher, J. Theor. Biol. 135, 455 (1988).
 R. M. Casey and A. S. Gaunt, J. Theor. Biol. 116, 45 (1985).
 I. R. Titze, Phonetica 28, 129 (1973) and 29, 1 (1974).
 D. H. Klatt and R. A. Stefanski, J. Acoust Soc. Am. 55, 822 (1974).
 J. H. Brackenbury, J. Theor. Biol. 81, 341 (1979).
 Lord Rayleigh, The Theory of Sound (Macmillan, London 1896; reprinted by Dover, New York, 1945) Vol. 2, Chapter 21.
 J. W. Coltman, J. Acoust. Soc. Am. 60, 725 (1976).
 N. H. Fletcher, Ann. Rev. Fluid Mech. 11, 123 (1979).
 R. C. Chanaud, Sci. Am. 222(1), 40 (1970).