

## Bird Song—A Quantitative Acoustic Model

N. H. FLETCHER†

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

*(Received 19 June 1987, and in revised form 2 May 1988)*

A detailed quantitative physical model is developed which gives a successful quantitative account of the “voiced” song of birds such as ravens. Using available anatomical and physiological data, the model allows calculation of syringeal membrane motion, volume flow waveform, tracheal pressure waveform, radiated acoustic power, and acoustic energy spectrum. Computed results for radiated acoustic power as a function of air-sac pressure and volume flow are in good agreement with measured values in the literature. The radiated power spectrum consists of exactly harmonic components at multiples of the vibration frequency of the syringeal membrane, the inharmonic modes of which are locked into frequency and phase coherence by the non-linearity of the driving force when the membrane strikes against the cartilage of the opposing air-passage wall. The spectral envelope has formant bands at the “closed-pipe” resonance frequencies of the trachea, supplemented by formant bands at slightly below the “open-pipe” resonance frequencies. The strengths and origins of these bands are made clear by the model. The computed power spectrum is in excellent agreement with the Sonagraph spectrum of *Corvus mellori* when anatomical parameters for this bird are used.

While the model is also able to produce inharmonic “screeched” song, attempts to produce a nearly pure-tone output by restricting the motion of the membrane so that it does not strike against the opposing cartilage have proved unsuccessful, the acoustic output being low in intensity and still exhibiting many overtones of the membrane frequency. This failure suggests that “whistled” song is produced in an entirely different manner from voiced song.

### 1. Introduction

The acoustics of bird song has been discussed in detail in the classic book by Greenewalt (1968), as well as in many other publications. Greenewalt provides references to the early literature while Brackenbury (1979*b*, 1982), Gaunt & Gaunt (1985) and Casey & Gaunt (1985) review more recent work. While some may consider that Greenewalt has provided an explanation of all the major features of the subject, it is clear that this view is not held by all workers in the area, and quite basic aspects of the acoustical processes involved continue to be the subject of speculation in the current literature (e.g. Casey & Gaunt, 1985). In the words of Gaunt & Gaunt (1985), however, much of the discussion is at the level of “informed supposition”.

As was clearly recognized by Greenewalt (1968) and re-emphasized by Gaunt & Gaunt (1985), there is no single simple mechanism used by all birds to produce

† Address for correspondence: Prof. N. H. Fletcher, Research School of Physical Sciences, Department of Electronic Materials Engineering, Australian National University, Canberra, ACT 2601, Australia

sounds for communication. Indeed consideration of human speech, in which voiced sounds, sibilants, fricatives, clicks and whistles are all employed by different cultures, should prepare us for a comparable variety of mechanisms in bird song.

It is not the purpose of the present paper to consider the whole gamut of these mechanisms in detail, but rather to concentrate on one of them. This concentration is not intended to minimize the importance of other mechanisms in particular cases, but simply serves to limit the field of discourse. In particular there will be no detailed discussion of sounds produced by purely aerodynamic means, without the aid of mechanically moving surfaces. Such aerodynamically produced sounds, which we may refer to generically as whistles, may well be important in some types of bird song (Gaunt *et al.*, 1982), and the acoustics of their production presumably involves the instability of air jets in interaction with apertures, edges and cavities. They may be discussed in analogy with the sound production mechanism in musical whistles, flutes and organ pipes (Rayleigh 1896; Powell, 1961; Chanaud, 1970; Coltman, 1976; Fletcher, 1979*a*), as has been recognized by Gaunt *et al.* (1982).

When there is mechanical motion of part of the vocal system in the course of sound production, an analogy can similarly be found with the mechanism of sound production in reed-driven woodwind musical instruments (Backus, 1961; Benade & Gans, 1968; Fletcher, 1979*a,b*), though the relative importance of source and resonator terms may be very different in the two cases. We shall refer to sounds produced in this way as "voiced", and the subsequent discussion in the present paper will be limited to such voiced sounds.

Both these sets of analogies are particularly fruitful because the study of sound production in musical instruments is much simpler than the biological case, and it has reached a stage of considerable quantitative sophistication. Understandably, however, this literature is little known to workers in the biological area, and there seems to have been little other than qualitative use made of the analogy. The fact that the behavior of musical instrument reeds or air jets is normally dominated by their associated tube resonator does not restrict the quantitative use of methods derived for musical instruments, provided account is taken of the different situation pertaining in vocal systems.

There are several distinct phases in the study of sound production in animals such as birds. In the first place the acoustical properties of the radiated sound (frequency spectrum, time evolution, power, etc.) must be known. A rich set of data on most of these features is given by Greenewalt (1968), and this has been supplemented by studies of typical radiated acoustic power levels by Brackenbury (1977, 1979*a*). Many further detailed studies of particular species can be found in the literature, but these compilations will be adequate for our present purposes. The next necessity is a knowledge of the anatomy of the sound-producing and sound-modifying organs. Again there is a considerable literature on this subject, but surveys in adequate detail for our present purpose have been given by Warner (1972), Brackenbury (1982), and Gaunt & Gaunt (1985). The final necessary observational input is a knowledge of relevant physiological variables for the system, such as air pressures and air flows. This information is less readily available, but some relevant data have been measured by Brackenbury (1977) and by Gaunt *et al.* (1982).

The next phase involves the postulation of a mechanism by which the observed sounds might reasonably be produced—a qualitative model. To be reasonably acceptable a model must be adequate in terms of physical principles and inherently plausible. It is possible for several competing models to provide equally reasonable explanations at this level, but sometimes crucial experiments can be devised to rule out some of the contenders. This is the level at which most discussion of avian vocalization mechanisms is carried out in the literature. Some particular models will be considered in the next section.

As a final test of any qualitative model, it is necessary to make it quantitative and to show that it is able to account for the observed acoustic output in terms of the known anatomical details and physiological variables. Specifically it must account for the acoustic power level, waveform and frequency spectrum of the song, using the observed values of internal air pressure and air flow. Only when we can construct a quantitative model with reasonable success can we claim to understand the sound production mechanism—until this stage is reached, we are still in the realm of “informed supposition”.

It is the purpose of the present paper to present such a quantitative model for voiced avian sounds. Before doing so, however, it is important to set out the philosophical basis on which it is constructed and the criteria that should be used in assessing its success.

The first principle is economy of hypothesis—Occam’s razor. The model should be as simple as possible, consistent with explaining the observed facts, and non-essential details of anatomy or physiology should be omitted. The model should also be free from undetermined parameters—all its quantities should be accessible (at least in principle) to physical measurement, though we may have to make do at this stage with reasonable estimates for some of them. The second principle is physical and mathematical consistency, which is a matter of physical insight and technical proficiency. As we shall see, the model is necessarily non-linear, so that no analytical solution is possible, but the approximations and assumptions made must be clear at every step.

Finally, since the anatomical and physiological input data to the model are all approximate, it is important to realize that we must be content with predictions which are in approximate agreement with observation. This means that the model must reproduce the main observed features of the acoustical output, and must give quantitative values which are correct to within a factor of order unity (i.e. between about 0.3 and 3). We shall return to this point when considering the behavior of our model.

## 2. Models for Voiced Song

The essential features of the avian vocal system are shown in Fig. 1, which is abstracted from the detailed anatomical work referred to above. Air from the lungs is expelled through the bronchi, trachea and mouth, passing on its way through the syrinx, or vocal organ, at the junction of the two bronchi with the trachea. One wall of each bronchus has a thin membranous section which lies nearly flat during normal

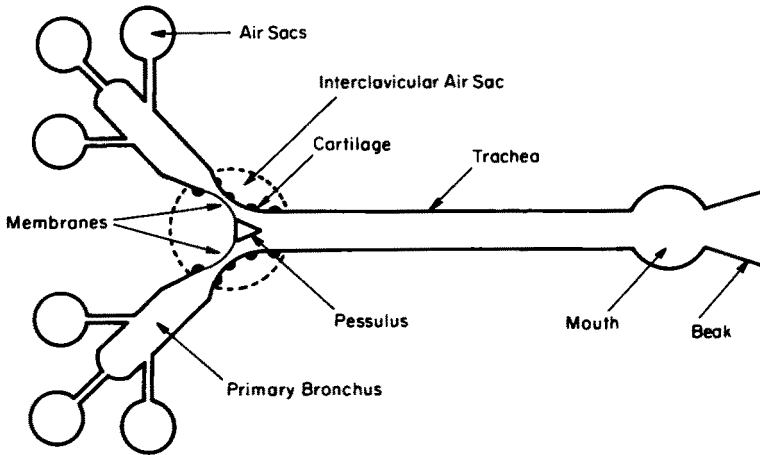


FIG. 1. Schematic view of the vocal system of a typical bird. Details of the implied anatomy are not important.

respiration and does not impede the air flow. In singing, the pressure in the interclavicular air sac is raised and the muscles of the syrinx are adjusted so that the membrane and its opposing cartilage protrude into the airway and interact with the air flow. It is this configuration that is shown in the figure. There are some variations in this basic anatomy (Gaunt & Gaunt 1985), but this model will suffice for our present purposes.

In voiced song the membranes are set into motion by the air flow and vibrate at a frequency which is determined by their mass and tension and by the nature of their interaction with the airstream. Some birds appear to be able to control the two membranes separately to produce two notes simultaneously (Greenewalt, 1968, pp. 55–78), but much song seems to employ only one source, or the two sources acting in unison. The radiated sound power is typically in the range 0.1 to 10 mW, but some birds can produce short calls with radiated acoustic power as high as 200 mW (Brackenbury, 1979*a*). Leaving aside those songs or parts of songs that are generally described as “whistled” and which have a nearly pure-tone character, most voiced sounds exhibit a series of harmonically related partials based on a fundamental whose frequency lies between about 100 Hz and 1 kHz for different species. The relative amplitudes of the partials are described by an envelope function which may emphasize one or more bands of frequencies, or formants. In song, the bird can adjust the fundamental frequency over a wide range, and may also embellish the sound with amplitude and/or frequency modulation at a rapid rate.

Several models have been proposed to describe the means by which the membranes of the syrinx are excited into motion and produce sound. Greenewalt (1968) has reviewed some of the earlier literature, and a definitive recent review has been given by Gaunt & Gaunt (1985). We will be concerned here only with recent models.

Greenewalt (1968) himself proposed that both “whistled” and voiced songs are produced by vibration of the membranes of the syrinx. In whistled sounds the

vibrational amplitude of the membrane is assumed to be small compared with its distance from the opposite wall of the trachea, so that its vibrations are nearly sinusoidal, giving an acoustic output with only a single frequency component. In voiced sounds, on the other hand, the vibration amplitude of the membrane is assumed to be large so that it approaches closely, or even contacts, the opposing tracheal wall, giving an inharmonic air flow with a multitude of harmonically related components. The frequency of vibration of the membrane is assumed to be controlled by its thickness and tension in a straightforward manner, while the motion itself is generated by Bernoulli forces in a manner which was not explored in detail. Greenewalt (1968) further showed that the trachea can act as an acoustic filter to emphasize or attenuate certain frequency ranges, the precise action depending on the relation between the acoustic impedance of the trachea and that of the syrinx. He found difficulty, however, in accounting for all features of the song spectrum in this way. Recent work by Nowicki (1987), who examined the change in spectral envelope of bird song in the transition between normal and helium-based atmosphere, confirms the importance of tracheal resonance in at least some bird songs.

More recently Brackenbury (1979*b*) has examined the aerodynamics of the sound production mechanism in greater detail, based upon a theoretical discussion, by Ffowcs Williams & Lovely (1975), of the behavior of a vibrating surface panel exposed to a uniform tangential air flow. This discussion was an attempt to add some quantitative detail to the general model of Greenewalt (1968), but in our view it did not succeed in this aim. The analysis of Ffowcs Williams & Lovely (1975) showed that a circular panel of radius  $a$  secured by a spring of strength  $K$  and exposed to an airflow of velocity  $v$  becomes unstable if  $A\rho av^2/K > 1$ , where  $\rho$  is the density of air and  $A$  is a constant of order unity. This effect, which is essentially the Bernoulli force, must certainly be taken into account in any model of the behavior of the syrinx, but modifications are necessary because the membrane is acted upon by a flow constrained to lie in a tube, rather than by a semi-infinite flow field.

Brackenbury (1979*b*) then went on to elaborate the model by assuming a non-linear term in the elasticity of the membrane in order to limit its static deflection. This is a reasonable assumption but is not necessary, as we shall see later. Finally he calculated the acoustic power radiated from the vibrating piston-like membrane, again using the results of Ffowcs Williams & Lovely (1975). The vibration amplitudes required to produce typical measured power levels were found to be unreasonably large, but the reason for this is clear. The model used refers to a piston radiating into a semi-infinite uniformly flowing medium, and the effect of the piston motion on the flow is small. In the syrinx, on the other hand, motion of the membrane constricts the windway significantly and has a major effect upon the flow. Clearly this effect must be included in any realistic model.

Finally we mention the model, or rather models, proposed by Casey & Gaunt (1985). These authors consider the pure-tone sounds of many bird songs, which we have described as whistles, to be produced by an entirely different mechanism from that used for voiced sounds. The mechanism proposed is a genuine whistle effect without the intervention of any vibrating surfaces. The production of such whistle

sounds can be described in terms of the unstable motion of jets of either varicose or sinuous form, or in terms of vortices (Chanaud, 1970), but an associated acoustic resonator is generally required to amplify the instability and stabilize the frequency. These details are not discussed by Casey & Gaunt (1985), but the mechanism is worthy of detailed investigation.

For the case of voiced song, Casey & Gaunt (1985) propose a variant of the mechanism put forward by Greenewalt (1968), with modifications to take account of what they perceive to be difficulties with this classic model. Specifically they express concern that the natural modes of a membrane with fixed boundaries do not constitute a harmonic series, and they propose this as a barrier to acceptance of the model as accounting for the harmonically related overtones observed in most voiced song. To overcome this problem they propose that the membrane is supported so that the two pairs of opposite boundaries (assuming the membrane to be rectangular) have different boundary conditions, so chosen that the resulting normal modes have harmonically related frequencies. This is their so-called "vibrating-string model" or VSM. There are two objections to the arguments of these authors. In the first place, as we shall show later, the spectrum of natural frequencies of the membrane is of relatively little importance for a non-linearly driven system. One dominant mode—usually the lowest mode—determines the vibrational frequency, all other modes are suppressed, and non-linear effects generate a precise and phase-locked series of harmonic overtones. There is therefore no need for concern about the classic model, at least on this ground. The second objection is more technical and refers to the mathematical assumptions underlying the VSM, which cannot in fact be realized as proposed. Since the VSM is, in any case, an unnecessary complication, it is not appropriate to discuss these objections in detail.

### 3. A Quantitative Acoustical Model

#### 3.1. ANATOMICAL MODEL

The simplified physical model upon which our acoustical model is based is shown in Fig. 2. We have simplified matters by considering only a single branch of the system, with a single syringeal membrane, but we shall return later to discuss the

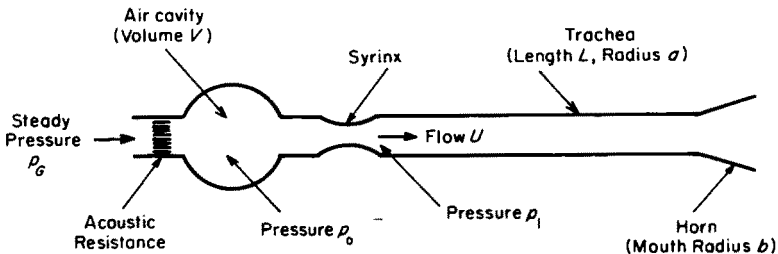


FIG. 2. Simplified acoustical model of the vocal system. The steady pressure  $p_G$  is the pressure in the air sacs which communicate with the bronchus through tubes with acoustic impedance  $Z_G$ . The volume  $V$  below the syrnix is essentially the volume of the primary bronchi.

behavior of the dual system. Other essential features of the physical model are as follows.

The air sacs in the body of the bird are assumed to contain air at a steady pressure  $p_G$  to provide the pneumatic power driving the whole system. All pressures mentioned in this connection are "gauge pressures" referred to normal atmospheric pressure as zero. The pressure unit is the pascal ( $100 \text{ Pa} = 1 \text{ cm water gauge}$ ). More technically, we have a pressure generator  $p_G$  in series with an acoustic impedance  $Z_G$  representing the effective flow resistance between the air sacs and the primary bronchus. This impedance is small compared with the impedance of the syrinx and trachea, as is demonstrated by low tracheal air pressures during normal respiration, and its precise value turns out not to be important.

The bronchus provides a small free volume  $V$  below the syrinx, while the trachea is a gently flaring tube leading to the mouth, which is assumed to be open for singing. The volume of the bronchus and dimensions of the trachea are readily estimated for anatomical studies, though these quantities clearly vary widely from one species to another. In the interests of simplification we replace the flaring trachea by a cylindrical pipe of the same average cross section. Since the change in diameter is small over the tracheal length, the acoustical effect of this substitution is small, but the true shape could be used in a more sophisticated model. We take the length of the trachea to be fixed, since only small adjustments seem possible and these will have little acoustic effect.

The geometry of the month is obviously complicated, and may change during singing in birds of some species. We shall return to consider this in more detail later, but the simplest assumption is to regard the month and beak as acoustically equivalent to a short flaring horn, the equivalent dimensions of which can be estimated with adequate accuracy for our purpose.

The syrinx itself is clearly the most critical part of the vocal system. During singing we assume it to have the geometry and dimensions illustrated in Fig. 3. A membrane of thickness  $d$  and width  $2h$  is assumed to protrude into the bronchus under the combined influence of muscular tension and the pressure in the external interclavicular air sac, as is described in more detail in Appendix A. During singing this external sac pressure will be largely counterbalanced by the internal pressure in the

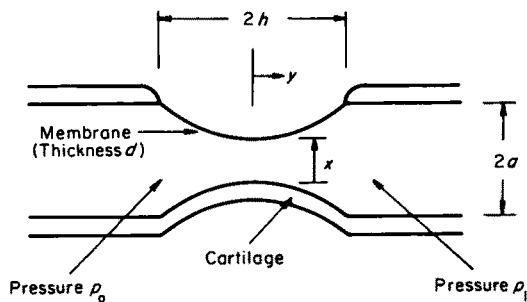


FIG. 3. Simplified model of the syrinx. Anatomical details are not important, but specific dimensions are required for the calculations of Appendix A.

bronchus, and this effect will be included automatically in our model. There is generally a ridge of cartilage on the opposing face of the bronchus. The shape and location of this ridge may be of some significance, as we discuss later, but for the present this refinement is not included in the model. All that is required for definiteness is that it should be possible for the membrane to touch the opposing cartilage in such a way as to close the air passage off completely. The co-ordinate  $x$  then measures the height of the free airway, assumed to extend completely across the bronchus. None of these details are essential, as we shall see presently, but we require that the model be explicit since we wish to write down explicit equations to describe it.

### 3.2. AIR FLOW THROUGH THE SYRINX

Now let us consider the volume flow  $U$  of air through the syrx. It is driven by the difference in pressure  $p_0 - p_1$ , where  $p_0$  is the pressure on the bronchial side of the constriction and  $p_1$  the pressure on the tracheal side. The bronchial pressure  $p_0$  is not equal to the pressure  $p_G$  supplied by the air sacs because of the finite impedance  $Z_G$  of the supply reservoir and the time variation in the flow. In fact the bronchus acts as an acoustic compliance of magnitude  $V/\rho c^2$ , where  $V$  is its volume,  $\rho$  is the density of air, and  $c$  is the velocity of sound in air (Beranek, 1954, pp. 128-129). The inflow is just  $(p_G - p_0)/Z_G$  and the outflow is  $U(t)$ , so that we can write

$$dp_0/dt = (\rho c^2/V)[(p_G - p_0)/Z_G - U]. \quad (1)$$

When the flow varies with time, as is the case during singing, we must allow for the force used to accelerate air through the syrx as well as for the steady flow (Backus, 1961; Fletcher, 1979*a,b*). This involves the acoustic inertance of the constriction (Beranek, 1954, pp. 131-132). We write therefore

$$p_0(t) - p_1(t) = CU^2 + D(dU/dt) \quad (2)$$

where, as is shown in the Appendix A,

$$C \approx \rho/8a^2x^2, \quad D \approx \rho/2(ax)^{1/2}. \quad (3)$$

These results are correct to within a factor of order unity for different syrx geometries. In using them, however, we must restrict  $x$  to be positive or zero.

### 3.3. MOTION OF THE SYRINGEAL MEMBRANE

The next part of the model involves description of the motion of the syringeal membrane under the influence of the aerodynamic forces in the constriction. It is a reasonable approximation to neglect tangential forces due to air viscosity and to consider only pressure forces, which include of course the influence of the Bernoulli term. On the upstream or bronchial half of the membrane the Bernoulli equation applies and we can write, for  $y < 0$ ,

$$p(y) = p_0 - \rho U^2/2S(y)^2 \quad (4)$$



where

$$S(y) \approx 2a[x + (a-x)(y/h)^2] \quad (5)$$

is the cross section of the airway at position  $y$  and we have omitted the small term corresponding to  $dU/dt$  in (2) in the interests of simplicity. Downstream from the constriction, where  $y > 0$ , it seems likely that the airflow forms a jet, so that the effective pressure acting on the downstream half of the membrane is just  $p_1$ .

To evaluate the effective force driving the fundamental mode of the membrane we must integrate the pressure, weighted by the membrane mode shape, over the entire membrane surface. Again, provided the membrane is nearly isometric, the result depends little upon shape and, as shown in Appendix A, has the magnitude

$$F(t) \approx 2A_1 ah \{ [(p_0 + p_1)/2] - \rho U^2 / [7(ax^3)^{1/2}] \} \quad (6)$$

where  $A_1$  is a constant of order unity. Because the flow-derived part of the force is concentrated near the center of the membrane where the airway is narrowest, it will tend to drive higher modes as well as the fundamental. The upstream-downstream pressure difference, on the other hand, has a broad distribution and will tend to drive most efficiently the lower modes, both symmetric and antisymmetric, of the membrane. As we shall see presently, it is the nature of the non-linear mechanism that one single mode will rapidly dominate the motion.

The shape and displacement law for the membrane are investigated briefly in Appendix A. Clearly the anatomy of the syrinx varies significantly from one bird species to another, and all that has been attempted is construction of a model for a sort of generic syrinx structure. To a good approximation, the syrinx membrane behaves like a simple taut membrane, irrespective of the fact that its shape is determined by the combined influence of membrane tension and air pressure in the interclavicular sac. This assumption has, indeed, also been made in all previous discussions of syrinx action.

The free vibration frequencies and damping coefficients of the membrane can be calculated once its dimensions, density  $\rho_M$  and effective tension  $T$  (which includes allowance for the pressure in the interclavicular sac) are known. Most important is the fundamental or lowest vibration frequency  $f_1$ , which we express more conveniently as the angular frequency  $\omega_1 = 2\pi f_1$ . For any nearly isometric membrane shape, we can see (Morse, 1948, pp. 172-208) that

$$T = (A_2/5)\rho_M ah d \omega_1^2 \quad (7)$$

where  $A_2$  is another constant of order unity. We shall use constants of order unity (by which we mean, of course, that they lie between about 0.3 and 3) frequently in our analysis to emphasize the fact that the results do not depend upon details of the assumptions of the model, such as the precise shape of the membrane. In performing illustrative calculations we shall set all these constants equal to unity for definiteness. The frequency of the second mode for such a nearly isometric membrane is about  $1.6\omega_1$  and that of the third mode about  $2\omega_1$ . Higher modes are even more closely spaced.

The behavior of the membrane itself can be described in terms of these normal modes. Suppose we concentrate on one mode, say the fundamental, and represent the central displacement associated with it by  $x_1$ . Other modes are treated in just the same manner. Then the equation of motion of that mode can be written as

$$m_1 \left[ \frac{d^2 x_1}{dt^2} + 2\kappa \frac{dx_1}{dt} + \omega_1^2 (x_1 - x_0) \right] = \varepsilon_1 F(t) \quad (8)$$

where  $\omega_1$  is the mode frequency,  $\kappa = \omega_1/2Q_1$  is the damping coefficient and  $x_0$  is the equilibrium position of the membrane in the absence of any pressure internal to the bronchus. The parameter  $\varepsilon_1$ , which we shall normally take equal to unity, measures the coupling between the forcing term  $F(t)$  and the mode in question, and depends both on the mode shape and on the position of the membrane in relation to the ridge of cartilage normally found on the opposite wall of the bronchial passage. The quantity  $m_1$  is an effective mass associated with the mode. It is less than the total mass of the membrane because we have chosen the coordinate  $x_1$  to measure the maximum displacement rather than the average displacement associated with the mode. To an adequate approximation we can write

$$m_1 = A_3 \rho_M \pi a h d / 4 \quad (9)$$

where  $A_3$  is a constant of order unity which differs slightly from one mode to another. The damping coefficient  $\kappa$  has its origin in viscous losses within the cellular material of the membrane itself, but these are supplemented by viscous losses within the flowing airstream. The actual value of  $\kappa$  for the syrxinx membrane is not known, but by analogy with other biological material it is probably not less than about 300 giving a  $Q$  value of about 10 at 1 kHz, or proportionally less at lower resonance frequencies. In the presence of the airstream the damping could well be increased (Ffowcs Williams & Lovely, 1975), and the  $Q$  value decreased, by a factor 2 or more.

As it stands, the left hand side of eqn (8) is linear, but there is reason for including at least some non-linearity to allow for the sharp change in behavior if the membrane strikes the opposite wall of the bronchus. This is most easily represented by requiring that

$$\kappa \rightarrow E\kappa \quad \text{if } x \leq 0. \quad (10)$$

The factor  $E$  is simply a large number of order 10 to 100 to allow for the damping effect of contact with the wall. Its real value is not known, since it depends on the "stickiness" of the contact between the membrane and the wall, but it is easy to vary this and see its effect.

It would be simple to add an elastic non-linearity of the type proposed by Brackenbury (1979*b*). This involves only the insertion of a factor of the form  $[1 + \gamma(x_1 - x_0)^2]$  in front of  $\omega_1^2$  in eqn (9). The coefficient  $\gamma$  then measures the strength of the non-linearity. It turns out, however, that inclusion of appreciable non-linearity of this type has other undesirable effects, so we shall omit it, or equivalently take  $\gamma = 0$ .

It is clear, however, that some non-linear effects must enter if the vibration amplitude of the membrane becomes large. The most likely such effect seems to be that the surrounding tissue, including the muscles tensioning the membrane, should

begin to participate in the motion. This would increase the effective vibrating mass  $m$  of eqn (9) in a way that can reasonably be described by an equation of the form

$$m \rightarrow m\{1 + \eta[(x - x_0)/h]^2\} \quad (11)$$

where  $\eta$  is a constant. If  $\eta \approx 10$ , the effect of the increase in mass will become significant when the displacement of the membrane is more than about one third of its radius, which is physically reasonable.

Clearly, for a complete treatment we should include a large number of possible vibrational modes of the membrane. This was the problem seen by Casey & Gaunt (1985). However the driving term  $F(t)$  is given from eqns (3)–(7) as a function of the displacement variable  $x$ , which is essentially just the sum of the mode displacements  $x_n$ . We should therefore really write

$$F(t) \rightarrow F(x_1, x_2, \dots). \quad (12)$$

The right hand side of eqn (12) does not depend explicitly upon time but is a complicated non-linear function of the variables  $x_n$  and their time derivatives. Such autonomous oscillating systems, albeit of much less complexity, have been extensively studied in the literature (Mickens, 1981). In particular, it has been shown (Fletcher, 1978) that, while independent excitation of more than one mode is possible, the usual result of strong non-linearity is that one mode becomes dominant and all others are effectively driven by it, resulting in an exactly harmonic and rigorously phase-locked vibration. This results in a complex but exactly repetitive waveform.

To show that this mode locking does in fact occur in our model system, we include just two modes with inharmonically related frequencies  $f_1$  and  $f_2$ . The addition of further modes of higher frequency is then largely irrelevant, since their behavior is similar.

#### 3.4. THE TRACHEA

Finally the model must describe the acoustic waves in the trachea—this makes no prior judgement about the importance of tracheal resonances. Suppose that a pulse of air with volume flow  $U(t)$  enters the trachea from the syrinx at time  $t$ . Then, as shown in detail by Schumacher (1981), this pulse creates a pressure wave  $Z_0 U(t)$ , where  $Z_0 = \rho c / \pi a^2$  is the characteristic impedance and  $a$  the radius of the trachea. Upon this pressure must be superposed, however, the pressures due to all previous flows into the trachea, which have been reflected from its open mouth end or from other irregularities. If the reflection behavior has the characteristic  $r(t)$ , then we find that

$$p(t) = Z_0 U(t) + \int_0^t r(t') [Z_0 U(t-t') + p(t-t')] dt'. \quad (13)$$

In the simplest possible case there is perfect reflection from the mouth after a time delay  $\tau$  which is just the round-trip transit time  $2L/c$  for a cylindrical pipe of length  $L$ . Because the termination is open there is a change in sign of the reflected wave, and we should also allow for some attenuation caused by wall losses, which

generally dominates losses by radiation from the mouth. We therefore write

$$r(t) = -\beta\delta(t - \tau) \quad (14)$$

where  $\delta(t - \tau)$  is the Dirac delta function, which is zero except when  $t = \tau$  and has unit integral. The coefficient  $\beta$  has the approximate value

$$\beta = 1 - 2\alpha L; \quad \alpha \approx 2 \times 10^{-5} \omega^{1/2} / a \quad (15)$$

where  $\alpha$  is the propagation attenuation coefficient (Benade, 1968).

In this approximation we have neglected the effect of mouth resonances on the reflection coefficient. It is simple to modify the form of  $r(t)$  to allow for these features, for example by replacing the delta function in eqn (14) by a sharply damped oscillation at the mouth resonance frequency, but at the present stage this only complicates the model unnecessarily. It turns out, however, to be desirable to broaden the delta function slightly to allow for increased attenuation in the trachea at high frequencies, according to eqn (15). We should also include an additive end-correction of about  $0.6b$ , where  $b$  is the radius of the effective horn-like mouth opening, in the length  $L$  assumed for the trachea.

### 3.5. SOUND RADIATION

Rather surprisingly, sound radiation does not need to be included in the primary dynamics of the vocal system, which is described to an adequate approximation by eqns (1)–(15). This is because, as we see later, the radiation of sound energy represents only a very small damping effect, accounting for only about 1% of the losses in the system. It is important, however, to show how the acoustic output can be calculated from the other variables.

The primary acoustic quantity inside the system is the pressure  $p_1(t)$  in the trachea near the syrinx. This variable is calculated as a function of time, and we need first to perform a Fourier transform to express it as a function of frequency,  $p_1(\omega)$ . This is easily performed numerically using a Fast Fourier Transform (FFT) algorithm. As shown in Appendix B, we can then calculate the power dissipated by  $p_1(\omega)$  in the radiation resistance at the mouth, transformed to the base of the tracheal tube. The result is then summed over all frequency components to give

$$P = \sum_{\omega} P(\omega) \approx \sum_{\omega} \frac{p_1(\omega)^2 R_R}{2Z_0^2 [\cos^2 kL + (\alpha L + R_R)^2 \sin^2 kL]} \quad (16)$$

where  $k = \omega/c$  and  $\alpha$  is the attenuation coefficient in the trachea, given by eqn (15).  $R_R$  is the resistive part of the radiation load at the mouth end of the trachea, after allowance has been made for the horn-like effect of the open beak, and is given approximately by

$$\begin{aligned} R_R &\approx Z_0 (kb/2)^2 && \text{for } kb < 2 \\ &\approx Z_0 && \text{for } kb > 2. \end{aligned} \quad (17)$$

The quantities  $P(\omega)$  give the power spectrum of the acoustic output and, when converted to acoustic levels  $L(f)$  in decibels by the familiar relation,

$$L(f) = 10 \log_{10} P(f), \quad f = \omega/2\pi \quad (18)$$

yield a plot of level as a function of frequency similar to that of a standard Sonagram section.

Because there is a phase change in each of the frequency components between the internal and external pressure waveforms, no simple waveform comparison can be made. If we require the radiated waveform, we must retain phase information in the FFT of the tracheal pressure and in the transformation to radiated pressure, and then perform an inverse FFT to recover the waveform.

#### 4. Numerical Evaluation of the Model

##### 4.1. INTRODUCTION

Equations (1)–(18), when supplemented by numerical values of the anatomical and physiological quantities involved, provide a complete mathematical model from which the behavior of the vocal system can be determined. Clearly the equations describing the system are complicated, and some of them are strongly non-linear. Experience shows that the non-linearity is an essential ingredient—if the equations are linearized by omitting the non-linear terms then the vibrations have indeterminate amplitude and the spectral behavior is quite different from that of the true system. There is therefore no alternative but to integrate the equations of motion numerically. Because of the particular way in which we have formulated the model, the numerical problem is relatively simple and the equations can be integrated on a micro-computer of modest power. This integration yields directly the waveforms of the membrane motion, the pressure in the bronchus and trachea, and the flow through the syrinx. From these quantities, as we have seen, it is possible then to calculate the power and the frequency spectrum of the acoustic radiation.

The values of anatomical and physical parameters used in an illustrative calculation are given in Table 1. They correspond to a bird of moderate size, such as a crow or raven, known to produce voiced song. None of the parameters are critical—they could be altered by as much as a factor 3 in either direction without affecting the qualitative behavior of the model, though clearly the quantitative behavior would change. This is in keeping with our general approach in which factors of order unity are neglected.

##### 4.2. INTERNAL VARIABLES

The general behavior of the model is summarized in Fig. 4, in which the major physical variables in the syrinx region are plotted as functions of time. Clearly the waveform of each variable is exactly repetitive, which indicates a complete frequency and phase locking of all the motions, and consequently exactly harmonic spectra, irrespective of the inharmonic relation of the normal mode frequencies of the syringeal membrane. In fact the mode of lowest frequency controls the motion because it is more compliant than higher modes, though this situation could be changed by altering the configuration of the membrane relative to the opposing cartilage so as to decrease the value of the coupling coefficient  $\epsilon_1$  for mode 1 in

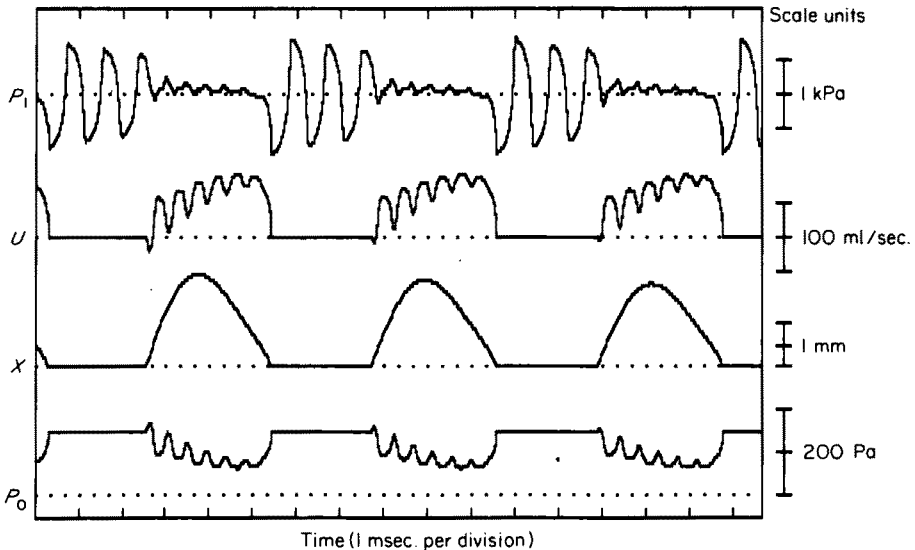


FIG. 4. Calculated waveforms for tracheal pressure ( $p_1$ ), volume flow ( $U$ ), membrane displacement ( $x$ ) and sub-syrinx pressure ( $p_0$ ), using the parameters shown in Table 1 and an air-sac pressure of 300 Pa.

eqn (8) and increase the corresponding value for mode 2. It is straightforward to describe how this could be achieved physically once the mode shapes are defined.

It is clear that the membrane, which we have taken to be in contact with the cartilage of the opposite wall of the bronchial tube in its equilibrium state, is forced open by the blowing pressure and then oscillates at a frequency somewhat greater than that of its fundamental mode under the combined influence of static and dynamic pressure forces. In the model the membrane closes against the wall once in each cycle of its motion, and we shall see presently that this feature is important. The amplitude of membrane motion depends on membrane thickness and blowing pressure, but is typically a few millimetres. There are obvious physical limits here, and a bird which uses a high blowing pressure cannot at the same time have a very thin membrane or it will be ruptured.

The pressure in the bronchus below the syrinx rises to equal the blowing pressure when the membrane closes the airway, but falls significantly during that part of the cycle during which the airway is open. The extent of this fall is determined by the impedance of the passages to the air sacs.

Turning to the acoustical variables  $U$ , the flow through the syrinx, and  $p_1$ , the pressure in the trachea just above the syrinx, we see the expected complex behavior. The flow is necessarily zero when the syrinx is closed, but is generally positive when the syrinx is open. During this latter part of the cycle, the flow exhibits pulsations because of reflections from the open mouth, the frequency of the pulsations being at essentially the "open-pipe" resonance frequency of the trachea, though lowered slightly by the loading of the constriction and sub-syrinx impedance.

The tracheal pressure shows similar small oscillations at its open-pipe frequency when the syrinx is open, but their amplitude is small because this pressure is being evaluated near one end of the pipe. When the syrinx is closed, on the other hand, the tracheal pressure shows oscillations at about half this frequency, corresponding to the "closed-pipe" frequency of the trachea. The amplitude of these oscillations is large because the pressure is being evaluated near the closed end of the pipe. These effects, it must be emphasized, arise naturally from the explicit physical assumptions of the model, set out above, and have not been imposed by any additional hidden assumptions.

It is also straightforward to evaluate power relations for the model. The first quantity of interest is the mean volume flow  $\bar{U}$ , which in this case has the reasonable value of 60 ml/sec. The pneumatic power expended by the bird in producing the vocal effect is now just the product of the air-sac pressure  $p_G$ , and the mean volume flow  $\bar{U}$ . For the case shown in the figure it amounts to about 18 mW. The power input to the syrinx is the time average of the product of the sub-syrinx pressure  $p_0(t)$  and the volume flow  $U(t)$ . In this case it amounts to about 11 mW, the balance of the original power having been expended in overcoming the flow resistance from the air sacs. This estimate is not unreasonable, but it would be reduced if the air-sac impedance were reduced. Finally the acoustic output from the syrinx is the time average of the product  $p_1(t)U(t)$ , which amounts to only about 2 mW.

#### 4.3. SOUND OUTPUT

Once we have the tracheal pressure waveform shown in the top line of Fig. 4, we can calculate the frequency distribution of the radiated acoustic power by making a Fourier transformation and applying eqns (16)–(18). Since the tracheal pressure waveform repeats at the membrane oscillation frequency, near 200 Hz in this case, the frequency spectrum will necessarily consist of a series of harmonics of this fundamental. Within each period, however, the pressure waveform shows that there is clearly a large amount of energy concentrated near the frequency of the closed-pipe resonances, in this case 1.2 kHz and its odd multiples. Finally, though the amount of pressure variation contributed by the lowered open-pipe resonances at multiples of 2.2 kHz is small, the transformation in eqn (17) increases the importance of these components.

The calculated power spectrum, shown in Fig. 5, clearly exhibits all these features. The whole spectrum consists of regularly spaced harmonic components based upon a fundamental near 200 Hz. There are broad formant bands (so called by analogy with bands produced in the same way in human speech) centered on the closed-pipe resonances 1.2, 3.6 and 6.0 kHz (marked A) and sharper formant bands (marked B) centered on 2.2, 4.4 and 6.6 kHz. These features of the spectrum are in excellent agreement with measurements reported by White (personal communication) for the Australian raven *Corvus mellori*. Details are given in Appendix C. In particular he remarks on the harmonic components and the formant bands corresponding to both closed and open-pipe tracheal resonances. The dimensions adopted in Table 1

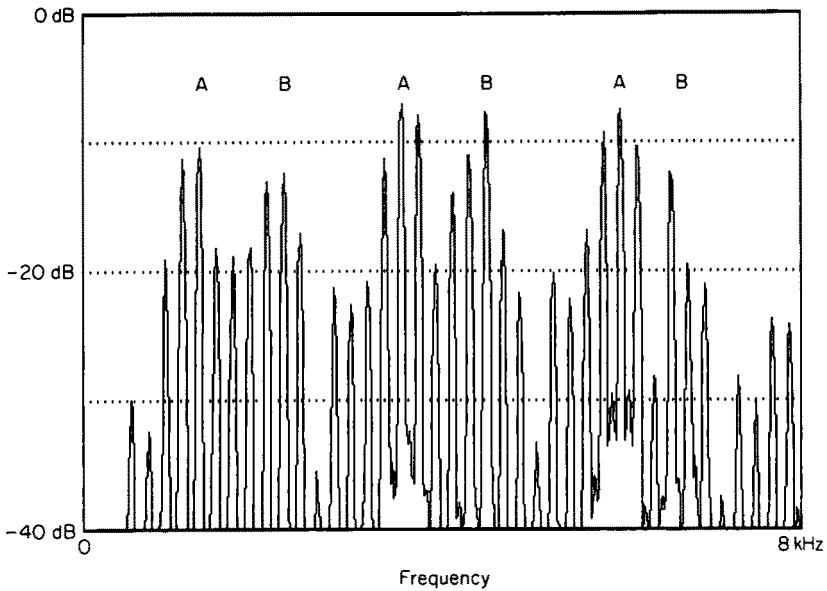


FIG. 5. Calculated power spectrum of the acoustic radiation from the model under the conditions of Fig. 4. The 0 dB reference is arbitrary, but the total radiated power is about 0.5 mW. Note the regularly spaced harmonics of the membrane vibration frequency, the "closed-pipe" tracheal formants (marked A) and the "open-pipe" tracheal formants (marked B). The open-pipe formant frequencies are lowered by the acoustic loading of the open syrxn.

TABLE 1  
*Assumed Values of Anatomical and Physiological Parameters*

Length of trachea	$L$	70 mm
Diameter of trachea	$2a$	7 mm
Diameter of membrane	$2h$	7 mm
Thickness of membrane	$d$	100 $\mu\text{m}$
Equilibrium opening <sup>†</sup>	$x_0$	0 mm
Membrane resonance frequencies <sup>†</sup>	$f_1$	150 Hz
	$f_2$	250 Hz
Membrane damping coefficient	$\kappa$	300 $\text{sec}^{-1}$
Membrane $Q$ at $f_1$	$Q$	2
Membrane nonlinear coefficient	$\eta$	10
Bronchial volume	$V$	1 ml
Effective mouth horn diameter	$2b$	20 mm
Blowing pressure <sup>†</sup>	$p_G$	0.3 kPa

(Quantities marked<sup>†</sup> are varied in the discussion)



correspond closely to those for one of the birds studied by White (personal communication). This perhaps represents the low-frequency limit for song fundamental, but the model is equally applicable to other cases if appropriate parameter values are used.

The 0 dB origin in Fig. 5 has no significance, but the total radiated power can be calculated by adding the contributions of all the components as specified in eqn (16). The result, for the specific case calculated, is about 0.5 mW. It is significant that this amounts to only about 1% of the total power expended by the bird in producing its song, a figure which is comparable with that reached by human players on musical wind instruments (Bouhuys, 1965), and in general agreement with the measurements of Brackenbury (1977, 1979*a*). The balance of the 2 mW of acoustic power output from the syrinx is expended in overcoming boundary-layer losses to the walls of the trachea.

It must be conceded, however, that the acoustic output of the model has an overemphasis on high-frequency components. This arises from the fact that all components of the vocal system were taken to be acoustically "hard" and the closure of the membrane against the opposing wall to be exact. Relaxation of both these assumptions, which were made in the interests of simplicity, would introduce a roll-off of high frequencies of at least 6 dB/octave, and perhaps as much as 12 dB/octave, above some cross-over point of a few kilohertz. This would improve the overall agreement between calculated and measured spectral envelopes.

#### 4.4. EXPLORATION OF THE MODEL

The model can be explored by varying its physical parameters within bounds limited by physical reality. Clearly the model can be scaled physically to represent larger or smaller birds and the properties of the syringeal membrane can be varied, but it is more interesting to leave the physical scale constant and to explore variations in vocal pitch and loudness. Leaving aside "whistled" song, it is known (Greenewalt, 1968) that many birds have a fundamental song frequency that can be varied over 2 to 3 octaves (a factor 4 to 8 in frequency)—a range very similar to that of an individual human singing voice. The model suggests this can be achieved by varying membrane tension and hence fundamental frequency, but we need to ensure that driving conditions remain reasonable over this range. Different species may, of course, have very different membrane thicknesses corresponding to different vocal ranges—and we must bear in mind that we need to know the thickness of the moist membrane before any drying out has occurred. The thickness chosen for our present version of the model probably represents a maximum for this size of bird. It is also interesting to calculate the range of acoustic output power available from a single version of the model when blowing pressure is varied.

In Fig. 6 only the blowing pressure is varied, other parameters of the model being held at the values shown in Table 1, and both air volume flow and radiated acoustic power are calculated. Clearly these two quantities both increase as blowing pressure is raised, and the values are reasonable in terms of the measurements of Brackenbury (1979*a*), when it is realized that the power output from the model also varies as

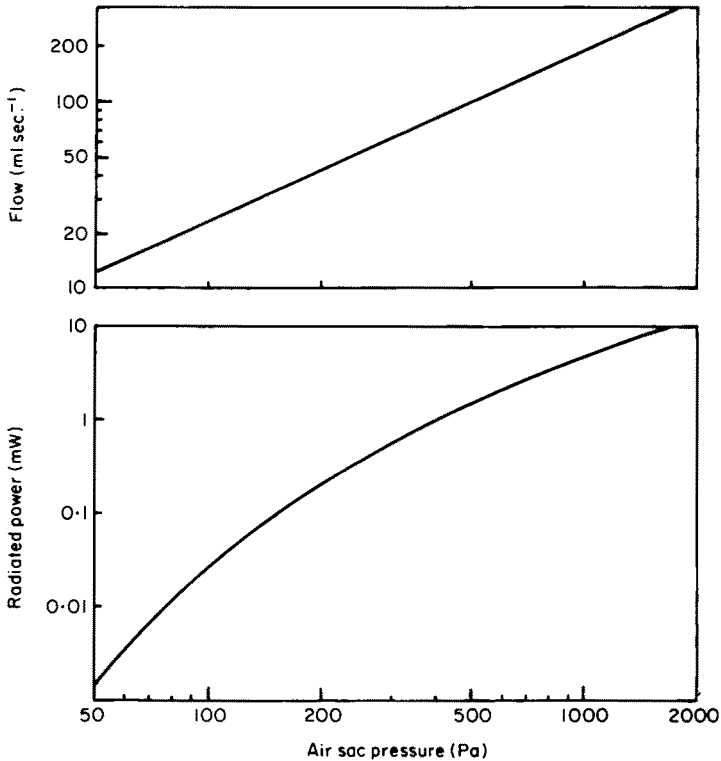


FIG. 6. Calculated radiated acoustic power and volume flow through the syrinx for the model parameters of Table 1, with the air-sac pressure varied between 50 and 2000 Pa (0.5–20 cm water gauge).

the square of the linear dimensions if the air sac pressure is constant. It is important to note that the fundamental frequency of the song produced by the model changes very little as the air-sac pressure is increased. This is a consequence of the fact that we have taken the membrane frequencies  $f_1$  and  $f_2$  to be constants. This assumption could be modified to account for the variation of fundamental frequency with loudness observed in some species (Gaunt & Gaunt, 1985), but construction of any realistic model would require much more specific information on the anatomy and muscular response of the system in question. The mean pressure in the trachea near the syrinx, incidentally, is about one tenth of the air sac pressure, though the r.m.s. pressure is very much higher.

Anatomical parameters for a domestic fowl of the species *Gallus domesticus* measured by Brackenbury (1977) are not immediately available, but it is clear that it is a good deal larger than the bird in our model. A peak acoustic power output of 200 mW for an air-sac pressure of 5–6 kPa and an air flow of 350 ml/sec is a reasonable order-of-magnitude extrapolation from the curve in Fig. 6, after a correction has been made for size.

Variation of the tension, and hence the resonance frequencies, of the syringeal membrane simply shifts the frequencies of the harmonic components shown in

Fig. 6 without changing greatly the formant envelope. The model produces sound over a fundamental frequency range up to around 2 kHz without difficulty, though the power output is reduced at the higher frequencies. It is also possible to use the model to account for shifts in formant frequencies when the bird sings in a helium-oxygen atmosphere, as in the experiments of Nowicki (1987), by simply changing the sound velocity  $c$  in the trachea. There seems, therefore, to be no difficulty in reproducing the main features of a typical voiced bird song.

#### 4.5. "WHISTLED" AND "SCREECHED" SONG

It is of interest to vary the operating parameters, as distinct from those specifying the anatomy, in an attempt to produce an acoustic output of a nearly sinusoidal character. If this could be achieved, then this would provide a reasonable model for the so-called whistled song, as suggested by Greenewalt (1968).

The obvious parameter to vary is the static position of the syringeal membrane relative to the opposing wall, as specified by the parameter  $x_0$ . Trials with this parameter set to give an equilibrium gap of 1 to 2 mm between the membrane and the wall showed a marked change of behavior when the blowing pressure was carefully adjusted so as to be insufficient to make the membrane meet the wall of the air passage, but still adequate to maintain it in motion. The whole system is then much more nearly linear, the sound output has only a few harmonics of the fundamental, and the radiated power level is quite low. However, because of the weakness of the non-linearity the second membrane mode is only partly suppressed, and frequency components arising from it are present in the song. It is difficult to regard this as even partial success in demonstrating the validity of the mechanism proposed by Greenewalt (1968, pp. 176–178) as the origin of "whistled" song. The bird clearly has detailed control over the anatomical and physiological parameters determining the behavior of its vocal system, and may be able therefore to produce a nearly pure-tone song more easily than has proved possible with the model. On the other hand, the low radiated sound power and the strength of the upper partials might well be taken to rule out this explanation and to give support to the view that this type of song is produced by an entirely different mechanism.

It is easier to achieve success in reproducing the raucous inharmonic screech characteristic of the song of birds such as cockatoos. For high air-sac pressure and high membrane tension the behavior of the model sometimes becomes non-periodic, quite irrespective of the existence of inharmonic modes on the membrane. In these circumstances the acoustic output consists of broad-band noise.

A word of caution should be said at this point about interpretation of all song showing harmonic components and formant bands as being necessarily voiced song of the type we have modelled. In particular, a song based on a single "whistled" tone, which is then subject to rapid and rather large-amplitude frequency modulation or non-sinusoidal amplitude modulation, will appear in a Sonagram as a set of equally spaced components grouped within a single band about the carrier frequency. Such a song could perhaps be produced by a "whistle" that is modulated by a

vibrating membrane, and the frequency of the supposed formant band would have no relation to any tracheal resonances.

#### 4.6. ELABORATIONS OF THE MODEL

While it is not reasonable to examine extensions of the model until its general correctness is accepted, it is appropriate to mention several elaborations or refinements that can reasonably be included and to outline the effects they could produce.

We might note in the first place that the trachea is not cylindrical, but increases in diameter from the syrinx to the mouth. It may also not be circular in cross section, but this has no significant acoustic effect provided the shape varies smoothly along the trachea. Treatment of wave reflections in flaring tubes is straightforward in principle but complex in practice (Schumacher, 1981; Morse, 1948, pp. 283–288). The overall acoustic effect is that the “closed-pipe” formant resonances are shifted upwards in frequency relative to the resonances of a cylindrical pipe of the same length, with the lower resonance frequencies increasing by a larger fraction than the upper ones. The “open-pipe” resonances, on the other hand, are unaffected by a simple tube flare. The fractional shifts for a flare rate typical of that of avian tracheae are quite small, however, and inclusion of this complication is not warranted. A flared trachea would, however, join more smoothly to the mouth horn, with a likely increase in radiative efficiency.

Similar remarks can be made about other modifications of the geometry of the trachea and mouth. Since the trachea is flexible, the bird may be able to stretch it to some extent by muscular tension. It seems most unlikely, however, that this stretching could exceed about 10% of the tracheal length, giving a similar uniform decrease in the frequencies of the formant bands. This may contribute to lack of constancy in the observed band frequencies during extended song, but it seems likely that other effects are more important.

The most effective way to modify tracheal resonance frequencies is by modification of the shape of the mouth. This is likely to be much less effective in birds than in mammals, however, for obvious anatomical reasons. In some birds with long hollow beaks, such as the Australian kookaburra *Dacelo gigas*, the mouth may, however, provide significant modification to the end-correction to the trachea and may possess cavity modes which affect different tracheal resonances differently. These effects would depend upon the extent of opening of the beak during singing, and could lead to significant relative shifts in formant frequencies. Subjective judgment of the song of the kookaburra suggests that these formant shifts do occur, but there does not appear to be any quantitative study in the literature. Such adjustments are, however, unlikely to be able to account for the very low formant frequencies observed in the vocal utterances of “talking birds” (Greenewalt, 1968, pp. 166–175), and some other explanation must be sought in this case (Gaunt & Gaunt, 1985).

We have remarked in passing that it is possible to excite either of the two modes assumed to be available to the syringeal membrane by adjustment of its position in relation to the opposing ridge of cartilage so as to make one of the coupling

coefficients  $\varepsilon_1$  of eqn (9), or its companion  $\varepsilon_2$  for the second mode, much smaller than the other. Our model shows this effect, which may well assist the bird in achieving sudden pitch changes in its song. Excitation of the second mode, which is approximately antisymmetric about the transverse axis of the membrane, essentially reproduces the "compression" or "ripple" model of Gaunt & Gaunt (1985). We should note, however, that the "travelling-wave" model that they suggest lacks, for the present, any physical basis.

Finally we return to the fact that our model differs from the actual bird in having only a single bronchus and syringeal membrane. It is important to enquire the effect of including the second sound source, and particularly the coupling to be expected between the two sources. It is, in fact, straightforward to generalize the model to include this second source, and the coupling between the two sources is readily seen to arise only from the fact that the tracheal pressure  $p_1$  acts on both membranes. They are otherwise independent. This coupling is actually quite weak, as we have seen for our single-source model, because the frequency of the tracheal pressure variations arising from pipe resonances is very much higher than that of the membrane motion. The only significant coupling will, in fact, come from the rather weak pressure components at the frequency of the membrane motion itself. This coupling may be adequate to lock the motion of the two sources in frequency and phase if they have been adjusted by the bird to nominal coincidence, but will impose little restriction on the independence of the two voices if their fundamental frequencies are well separated (Fletcher, 1978).

It is interesting to speculate that the bird may in fact adjust the membrane configuration of its two voices so that a different mode is excited on each, giving a considerable frequency separation without a great imbalance of muscular tension.

## 5. Conclusions

Reviewing the previous section, it is clear that the model we have presented is able to simulate many aspects of the observed and measured vocal achievements of at least certain species of birds. In particular, it shows that the radiated power spectrum of voiced song will generally exhibit a large number of precise harmonics of a low fundamental frequency associated with the motion of the syringeal membrane, irrespective of the fact that the natural modes of that membrane may themselves be very far from having a harmonic relationship. In addition, it gives a good account of the nature and frequencies of the formant bands reported by Greenewalt (1968) and many other workers, and produces quite detailed agreement with the acoustical and anatomical results reported by White (1988). It shows in detail how these formant bands arise, and demonstrates their relation to the "open-pipe" and "closed-pipe" resonances of the trachea.

More than this, the model produces quantitative predictions for the radiated acoustic power and air flow through the vocal system which are in acceptably good agreement with the measurements of Brackenbury (1977, 1979*a*). It is not meaningful to vary the parameters of the model in search of precise agreement in the absence of detailed anatomical and physiological data with which to compare the values of

the parameters used to achieve best fit. The model does, however, allow these parameters to be inserted, when they become available from future research, in order to test the level of detailed prediction achieved.

By varying the parameters of the model within reasonable bounds, it has also proved possible to reproduce the wide-band non-harmonic "screeched" sound made by some birds, confirming that we have achieved a reasonable understanding of this mode of utterance as well. Efforts to produce nearly pure-tone sounds, generally described as "whistles", have been unsuccessful, though a rather low-intensity song with much less harmonic development is certainly produced if the configuration of the syrinx is adjusted so that the vibrating membrane is prevented from contacting the opposite wall of the airway. While this does not completely rule out this mechanism for the production of "whistled" sounds, it does suggest that detailed consideration should be given to other possibilities, such as those proposed by Gaunt *et al.* (1982).

In summary, it can be claimed that this model places on a firm foundation of physical understanding, for the first time, one of the prime mechanisms of avian vocalization. It is hoped that, in due course, other mechanisms that have been suggested for different types of avian sound production can be subjected to similar quantitative analysis, for it is only by such analysis that such hypotheses can be tested against reality.

I am most grateful to Sir Frederick White for bringing this interesting problem to my attention, and for educating me in aspects of ornithology. My thanks are also due to one of the reviewers for this Journal who pointed out several significant pieces of work in the literature, of whose existence I was ignorant.

#### REFERENCES

- BACKUS, J. (1961). Vibrations of the reed and the air column in the clarinet. *J. acoust. Soc. Am.* **33**, pp. 806-809.
- BENADE, A. H. (1968). On the propagation of sound waves in a cylindrical conduit. *J. acoust. Soc. Am.* **44**, pp. 616-623.
- BENADE, A. H. & GANS, D. J. (1968). Sound production in wind instruments. *Ann. N.Y. Acad. Sci.* **155**, pp. 247-263.
- BERANEK, L. L. (1954). *Acoustics*, reprinted 1986, pp. 128-129, 131-132. New York: American Institute of Physics.
- BOUHUYS, A. (1965). Sound-power production in wind instruments. *J. acoust. Soc. Am.* **17**, pp. 453-456.
- BRACKENBURY, J. H. (1977). Physiological energetics of cock crow. *Nature, Lond.* **270**, pp. 255-259.
- BRACKENBURY, J. H. (1979a). Power capabilities of the avian sound producing system. *J. Expl. Biol.* **78**, pp. 163-166.
- BRACKENBURY, J. H., (1979b). Aeroacoustics of the vocal organs of birds. *J. theor. Biol.* **81**, pp. 341-349.
- BRACKENBURY, J. H. (1982). The structural basis of voice production and its relationship to sound characteristics. in: *Acoustic Communication in Birds* (D. E. Kroodsma & E. H. Miller, eds) Vol. 1, pp. 53-73. New York: Academic Press.
- CASEY, R. M. & GAUNT, A. S. (1985). Theoretical models of the avian syrinx. *J. theor. Biol.* **116**, pp. 45-64.
- CHANAUD, R. C. (1970). Aerodynamic whistles. *Scient. Am.* **222** (1), pp. 40-46.
- COLTMAN, J. W. (1976). Jet drive mechanisms in edge tones and organ pipes. *J. acoust. Soc. Am.* **60**, pp. 725-733.
- FFOWCS WILLIAMS, J. E. & LOVELY, D. J. (1975). Sound radiation into uniformly flowing fluid by compact surface vibration. *J. Fluid Mech.* **71**, pp. 689-700.
- FLETCHER, N. H. (1978). Mode locking in non-linearity excited inharmonic musical oscillators. *J. acoust. Soc. Am.* **64**, pp. 1566-1569.

- FLETCHER, N. H. (1979a). Air flow and sound generation in musical wind instruments. *A. Rev. Fluid mech.* **11**, pp. 123–146.
- FLETCHER, N. H. (1979b). Excitation mechanisms in woodwind and brass instruments. *Acustica* **43**, pp. 63–72.
- GAUNT, A. S. & GAUNT, S. L. L. (1985). Syringeal structure and avian phonation. *Curr. Ornithology*, **2**, pp. 213–245.
- GAUNT, A. S., GAUNT, S. L. L. & CASEY, R. M. (1982). Syringeal mechanics reassessed: Evidence from *Streptopellia*. *Auk* **99**, pp. 474–494.
- GREENEWALT, C. H. (1968). *Bird Song: Acoustics and Physiology*. pp. 55–78, 176–178, 166–175. Washington: Smithsonian Institution Press.
- MICKENS, R. E. (1981). *An Introduction to Nonlinear Oscillations*. New York: Cambridge University Press.
- MORSE, P. M. (1948). *Vibration and Sound*. reprinted 1976, pp. 172–208, 283–285. New York: American Institute of Physics.
- NOWICKI, S. (1987). Vocal tract resonances in oscine bird sound production: Evidence from birdsongs in a helium atmosphere. *Nature, Lond.* **325**, pp. 53–55.
- OLSON, H. F. (1957). *Acoustical Engineering*. pp. 88–112. Princeton: Van Nostrand.
- POWELL, A. (1961). On the edgetone. *J. acoust. Soc. Am.* **33**, pp. 395–409.
- RAYLEIGH, J. W. S. (1896). *The Theory of Sound*, reprinted 1945, Vol. 2, pp. 376–478. New York: Dover.
- SCHUMACHER, R. T. (1981). *Ab initio* calculations of the oscillations of a clarinet. *Acustica* **48**, pp. 71–85.
- WARNER, R. W. (1972). The anatomy of the syrinx in passerine birds. *J. Zool. Lond.* **168**, pp. 381–393.

## APPENDIX A

Referring to Fig. 3, we approximate the shape of the membrane by a quadratic form, so that the height of the opening is

$$z = x + (2a - x)(y/h)^2 \quad (\text{A1})$$

where  $y$  is measured along the length of the tube. We assume the opening between the membrane and opposing cartilage to be approximately rectangular, with a width  $2a$  when viewed along the tube axis.

Consider a state in which there is a pressure  $p_G$  in the interclavicular sac and pressure  $p_0$  in the interior of the syrinx. The membrane will then have a smooth curvature as shown, and we may take the curvature radius to be  $R$  (the two principal curvatures may in fact be different, but we neglect this complication as unimportant). The height of the bulge in the membrane is  $b = 2a - x - z$ , where  $z$  is the height of the cartilage. Now from simple geometry

$$b(2R - b) = h^2 \quad (\text{A2})$$

and pressure equilibrium requires that

$$p_G - p_0 = AT/R \quad (\text{A3})$$

where  $T$  is the tension in the membrane and  $A$  is a constant of order unity. Provided  $b \ll 2R$ , then

$$b \approx h^2(p_G - p_0)/AT. \quad (\text{A4})$$

The membrane tension will have two components, a static component supplied by muscle tension, and a dynamic component caused by the slight stretching of the membrane as the distance  $b$  changes. If the muscle tension component is dominant, as is very likely, then eqn (A4) shows that the deflection  $b$ , or equivalently  $x$ , varies

linearly with the tracheal pressure  $p_0$ . The membrane can therefore be treated in essentially the same way as a simple taut membrane. For convenience we take as a reference state one in which the tracheal pressure  $p_0 = 0$  and the displacement coordinate  $x = x_0$ .

If the general case, suppose the volume flow in the tube is  $U$ , then the flow velocity at position  $y$  is  $U/2az$ . By Bernoulli's theorem, the quantity  $p + \rho v^2/2$  is constant in the flow, so that if  $p_0$  is the static pressure in the bronchus, then the static pressure at the point  $y$  where the height of the opening is  $z$  is

$$p(y) = p_0 + \frac{1}{2}\rho U^2[(1/\pi a^2)^2 - (1/2az)^2] \approx p_0 - \rho U^2/8a^2z^2 \quad (\text{A5})$$

where the last form of writing applies if  $z \ll a$ .

The force acting on the upstream half of the membrane can be evaluated by integrating  $p(y)$  over this part of its area. If the flow is assumed to separate from the membrane downstream from the constriction, then the pressure over this half of the membrane is simply  $p_1$ , the pressure at the base of the trachea. The integral is a little complicated,

$$F \approx \frac{\rho U^2 h}{4x^2(2a-x)} \left[ \left( \frac{x}{2a-x} \right)^{1/2} \tan^{-1} \left( \frac{2a-x}{x} \right)^{1/2} + \frac{x}{2a} \right] \quad (\text{A6})$$

but it reduces to the form given in eqn (3) of the main text if  $x \ll 2a$ . If this condition is not fulfilled, then the dynamic pressure is small and the total expression is still approximately correct.

The acoustic inertance of the constriction is evaluated similarly as

$$D \approx \int_{-h}^h \frac{\rho dy}{2az}. \quad (\text{A7})$$

Evaluation is straightforward and the result

$$D \approx \frac{\rho h}{2ax} \left( \frac{x}{2a-x} \right)^{1/2} \tan^{-1} \left( \frac{2a-x}{x} \right)^{1/2} \quad (\text{A8})$$

can be reduced to that given in eqn (3) of the main text in the limit  $x \ll a$  or can be ignored if this condition is not fulfilled.

## APPENDIX B

The trachea is a lossy tube of length  $L$ , terminated with a radiation impedance

$$Z_R = R_R + jX_R \quad (\text{B1})$$

where, as usual,  $j = -\sqrt{-1}$ . We can incorporate the reactive part  $jX_R$  as a correction to the length  $L$ , and we then find (Olson, 1957) that the impedance presented at the syrxinx end of the trachea at frequency  $\omega$  is

$$Z(\omega) = Z_0 \left[ \frac{R_R \cos kL + jZ_0 \sin kL}{jR_R \sin kL + Z_0 \cos kL} \right] \quad (\text{B2})$$



where  $Z_0 = \rho c / \pi a^2$  is the characteristic impedance of the tracheal tube. Because the trachea has viscous and thermal losses at its walls, expressed by the attenuation coefficient  $\alpha$  given by eqn (15) of the main text, the propagation constant  $k$  is complex and is given by

$$k = \omega / c - j\alpha. \quad (\text{B3})$$

If we express the pressure  $p_1(t)$  at the base of the trachea in the frequency domain by its Fourier transform  $p_1(\omega)$ , and similarly express the radiation resistance  $R_R(\omega)$  as a function of frequency, then the power dissipated in the radiation resistance, and hence the radiated acoustic power, is given by

$$P = \frac{1}{2} \int_0^\infty p_1(\omega)^2 \operatorname{Re} [Z(\omega)^{-1}] d\omega \quad (\text{B4})$$

where  $\operatorname{Re}[Z^{-1}]$  is the real part of the admittance  $Z^{-1}$ . Using eqns (B2) and (B3), we find

$$P \approx \frac{1}{2} \int_0^\infty \frac{p_1(\omega)^2 R_R(\omega)}{Z_0^2 \sin^2(\omega L / c) + [R_0^2 + Z_0^2(\alpha L)^2] \cos^2(\omega L / c)} d\omega. \quad (\text{B5})$$

If we use a discrete rather than a continuous Fourier transform, this integral should be re-expressed as a sum over the spectral components.

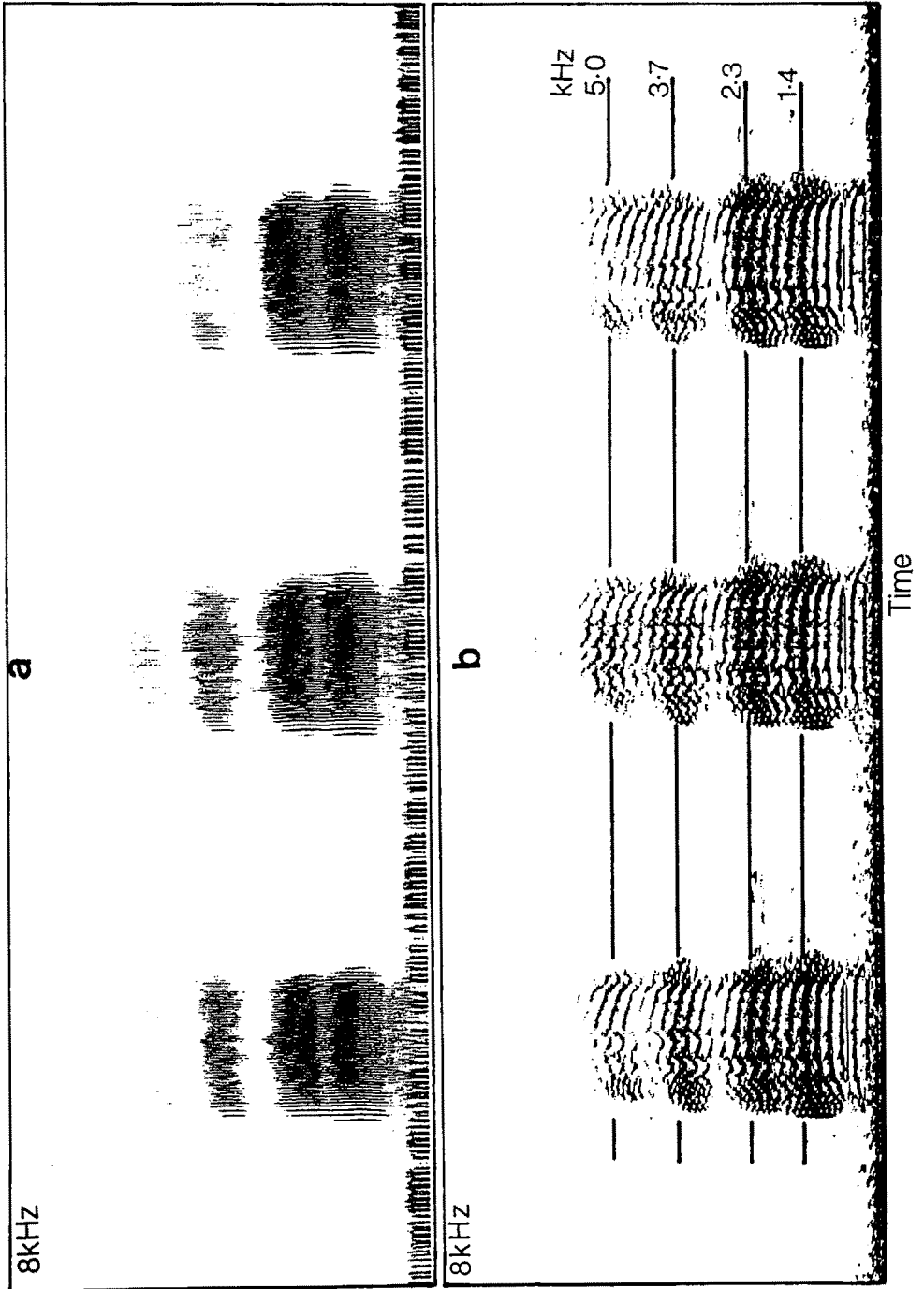
Finally we need to know the form of  $R_R(\omega)$ , allowing for the effect of the mouth horn, which increases the effective radius of the trachea from  $a$  to  $b$  over a short distance. The transformation is somewhat complex (Olson, 1957) and depends upon the shape of the horn. The mouth horn has very little effect on the total radiated power at frequencies for which it is short compared with the wavelength; it serves rather to increase the directionality of the radiation. When its length is an appreciable fraction of a wavelength, however, it acts as an acoustic transformer and raises the resistive part of the impedance as its throat. It is an adequate approximation for our present purpose to take

$$\begin{aligned} R_R(\omega) &\approx (\rho c / \pi a^2)(kb/2)^2 & \text{for } kb < 2 \\ &\approx (\rho c / \pi a^2) & \text{for } kb > 2 \end{aligned} \quad (\text{B6})$$

where  $k = \omega / c$ . This approximation overestimates the power radiated at low frequencies, but the error is not important here.

#### APPENDIX C

White (personal communication) has made an acoustical and anatomical study of several species of Australian ravens. These birds produce a simple cry, consisting of short sounds repeated about once a second. The repetition rate is regular and the individual sound segments are nearly identical. A Sonograph analysis of the cry of the little raven *Corvus mellori*, reported by White, is shown in Fig. 7. The upper part of the figure uses a wide-band filter to optimise time resolution, while the lower part uses a narrow-band filter to optimise frequency resolution. Fig. 7(b) shows that



the song segments have a frequency spectrum consisting of harmonics based upon a fundamental of frequency about 200 Hz. This fundamental frequency is approximately constant but rises at the beginning of each segment and falls at the end. These features are also apparent in Fig. 7(a), in which the separate air pulses through the vocal tract are resolved.

Figure 7 also shows that the song has a series of formant bands with center frequencies around 1.4, 2.3, 3.7 and 5.0 kHz. These frequencies are close to integral multiples of a fundamental of about 1.3 kHz, though the deviations from this simple relationship appear to be real.

Anatomical study of an individual of the species, which was of average size, showed that its trachea is about 70 mm in length and elliptical in cross section, with an effective diameter which increases from about 6 mm at the syrinx to about 8 mm at the mouth. If the trachea were to be approximated by a cylindrical tube, then its "closed-pipe" resonances, corresponding to a pressure maximum at the syrinx, would be odd multiples of a fundamental of about 1.2 kHz, and its "open-pipe" resonances even multiples of the same fundamental. The effect of the slight flare in the trachea is to raise the frequencies of the odd resonances by a small amount, while leaving the even resonances unaffected. These frequencies, simplistically calculated, are thus in quite good agreement with those of the measured formant bands. More importantly, the whole spectrum of the song is in very good agreement with the spectrum calculated from our detailed model, as can be seen from a comparison of Fig. 5 with Fig. 7(b). As well as being slightly influenced by tracheal taper, the resonance frequencies are modified by matching conditions below the syrinx.

White also studied several other species of corvids. They all showed similar call spectra, with harmonic components and formant bands. All the other species studied, however, had voice fundamental higher in frequency than that of *C. mellori*, even though some of them were larger birds, which made the clear identification of formant bands in their song less striking.

---

FIG. 7. Sonagraph analysis of the song of the Australian raven *Corvus mellori* (White, personal communication) using (a) a wide-band filter for optimal time resolution, and (b) a narrow-band filter for optimal frequency resolution. The record length is 2.2 sec, and the frequency span is 0 to 8 kHz in each case. The spectrum consists of harmonics of the syrinx frequency (about 200 Hz) with a formant envelope determined by vocal-tract resonances.