How Do Birds Sing?
Author(s): Stephen Nowicki and Peter Marler
Reviewed work(s):
Source: *Music Perception: An Interdisciplinary Journal*, Vol. 5, No. 4, Biological Studies of Musical Processing (Summer, 1988), pp. 391-426
Published by: University of California Press
Stable URL: http://www.jstor.org/stable/40285408
Accessed: 01/07/2012 16:19

Music Perception
Summer 1988, Vol. 5, No. 4, 391–426

How Do Birds Sing?

STEPHEN NOWICKI & PETER MARLER
The Rockefeller University

Hypotheses are reviewed about how the vocal apparatus of birds operates during singing, focusing especially on the means by which the tonal sounds so typical of birdsong are generated. Evidence for the widely held view that the "two voices" in the songbird syrinx act independently of one another, and independently of acoustic resonances of the vocal tract, is found to be incomplete. Results of several experiments, including effects of placing singing birds in helium air, indicate that coupling of the "two voices" can occur and that changes in vocal tract resonances modify the tonal quality of birdsongs. A new model of songbird phonation is proposed, implying close coordination between the operation of the acoustic sources in the syrinx and dynamic modifications of the vocal tract.

Introduction

As scientists studying animal behavior, we are often so engrossed by the functional and evolutionary significance of birdsong as a communication signal that we forget the powerful aesthetic feelings it holds for us as a kind of natural music. Such was not the case for an earlier generation of scientists. For those who lacked the technology for recording bird sounds and for analyzing their structure with devices such as the sonagraph (Potter, Kopp, & Green, 1947), it was common practice to annotate birdsongs on a musical staff (e.g., Mathews, 1921; Figure 1). Some ethologists even suggested that birdsong must be viewed as primitive art, beautiful from the bird's point of view as well as our own (Mathews, 1921; Saunders, 1929; Craig, 1943; Koehler, 1951; Hartshorne, 1958; Thorpe, 1961; Armstrong, 1963; Hall-Craggs, 1969). As a prelude to his discussion of "vocal and instrumental music" in birds, Darwin (1874, p. 697) remarked: "On the whole, birds appear to be the most aesthetic of all animals, excepting of course man, and they have nearly the same taste for the beautiful as we have."

Requests for reprints may be addressed to either author at The Rockefeller University Field Research Center, RR #2, Box 38B, Tyrrel Road, Millbrook, New York 12545.

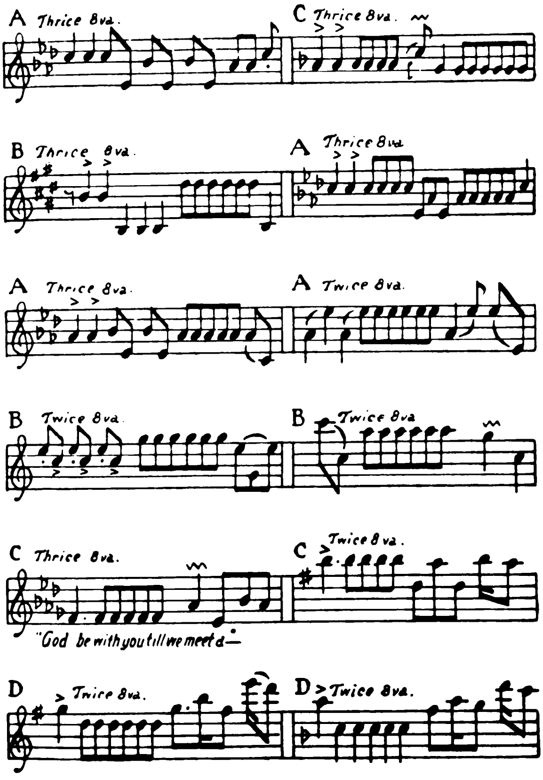



Fig. 1. Song sparrow songs transcribed using Western musical notation. Shown are 12 songs from four individuals (A, B, C, and D). Note that songs differ from bird to bird and that each individual has several song types in his repertoire. From Mathews (1921, p. 113), reprinted by permission of Dover Publications, Inc., New York.

If the sounds of birds share characteristics with human music, what of the instrument with which these sounds are produced? Here, too, birds are unique. Unlike mammals, the bird's larynx is presumed to play little or no role in sound production. Instead, the principal vocal organ of all birds is the syrinx, an organ found in no other animal group. The anatomy of the syrinx, its position relative to the rest of the vocal tract, and its variously proposed modes of operation all differ radically from what is known about the larynx.

The musicality of bird sounds, and the uniqueness of the structures responsible for their production, have led many to propose analogies between the functioning of the syrinx and the production of sound by certain musical instruments. In one of the earliest papers on the subject, Hérissant

(1753) concluded that the vocal organ of birds acts much the same as the double reed of an oboe. The famed anatomist Georges Cuvier (1800, 1805) agreed with the instrumental analogy, but offered the French horn and the trombone as more accurate comparisons. Similarly, organ pipes and other reed or brass instruments have been suggested as models for avian vocal production (e.g., Häcker, 1900; Réthi, 1908; Rüppell, 1933).

Each analogy with a musical instrument implies a conception of how sound is produced by the bird's vocal system. For example, the acoustic source of an oboe, a vibrating double reed, is quite different in its physical and acoustic characteristics from those of vibrating human lips, which form the acoustic source of a trombone. Thus, the differing views of Hérissant and Cuvier reflected their contrasting interpretations of how vibration is induced in the syrinx. One feature is shared, however, by all such analogies. In woodwind and brass instruments, the vibration of the source is strongly influenced by the acoustical properties of the air column with which it is associated. That is, the source is more or less constrained to produce acoustic oscillations at frequencies corresponding to resonances of the attached tube (Benade, 1976). These instrumental analogies suggest, then, that changes in the resonance characteristics of a bird's vocal tract assume paramount importance in determining the fundamental frequency of the song it produces.

Other researchers have suggested that, in spite of anatomical differences, syringeal operation may be analogized to human phonation (Thorpe, 1959; Klatt & Stefanski, 1974). The most significant implication of the human model is the functional relationship between the acoustic source on the one hand and associated acoustic resonances of the rest of the vocal tract on the other. In human phonation, the oscillation of the vocal folds is generally uninfluenced by the acoustic properties of the vocal tract. Instead, the tract resonances act as an acoustic filter, selectively attenuating some frequencies while letting others pass freely. As the human vocal tract changes shape during speech, there are shifts in these resonances (known as formats), creating differing amplitude spectra in the emitted sounds (Fant, 1960).

In more recent times, analogies of syringeal function with either musical instruments or human speech have fallen into disfavor. With the advent of new technology for sound analysis, some general acoustic properties of birdsong became more apparent, greatly complicating comparisons with other sound producing systems. The most outstanding of these properties is the presence in many birdsongs of two tones, not harmonically related, that are independently modulated. Early anatomical investigations suggested that the songbird syrinx could be a functionally double vocal organ (Cuvier, 1800; Savart, 1826; Häcker, 1900). However, it awaited the unambiguous documentation of the presence of two sounds not readily attributable to a single source, with the sound spectrograph, for the "two-voice theory" to become widely accepted (Potter et al., 1947; Borror & Reese, 1956; Greenewalt, 1968).

The two-voice theory seemed to discount the possibility that acoustic resonances have any influence over the vocal output of a bird. Two sources could not act independently if their behavior was determined by the acoustic properties of a single tube (Greenewalt, 1968). At the same time, detailed models of how oscillation is induced in the syrinx suggested that the acoustic sources are simple vibrating membranes and that most if not all of the acoustic properties of birdsong, including frequency and amplitude modulations, could be accounted for by the independent action of two membrances (Greenewalt, 1968; Stein, 1968; Gaunt & Wells, 1973). Thus, for the last 20 years the most widely accepted theory of songbird phonation has held that the syrinx possesses two functionally distinct and independent acoustic sources, that the action of these sources alone is responsible for the acoustic quality and variations observed in song and that acoustic resonances of the vocal tract above the syrinx play no role in determining or modifying sound quality. According to this view, the earlier analogies to instruments and voice seem inappropriate and counterproductive.

We recently began to reassess some of the underlying assumptions about how birds make sound, especially with regard to the two-voice theory and the involvement of acoustic resonances in production. New experimental findings suggest that current models of avian phonation need to be revised yet again. Further, we find ourselves returning to older analogies with wind instruments and the human voice as sources of insight into the mechanisms of avian phonation. In this article, we briefly review theories of phonation and then point out where recent results suggest that a new view must be taken. We begin with an overview of the kinds of sounds found in birdsong and end with a new model of how birds sing. Because syringeal structure and its presumed function differs among different taxonomic groupings of birds, we will confine our remarks to just one group, the Oscines or "true songbirds." This choice is appropriate not only because members of this group are among the most versatile vocal performers, but also because it is the most numerous avian taxon, including close to half of all living species of birds (Welty, 1982).

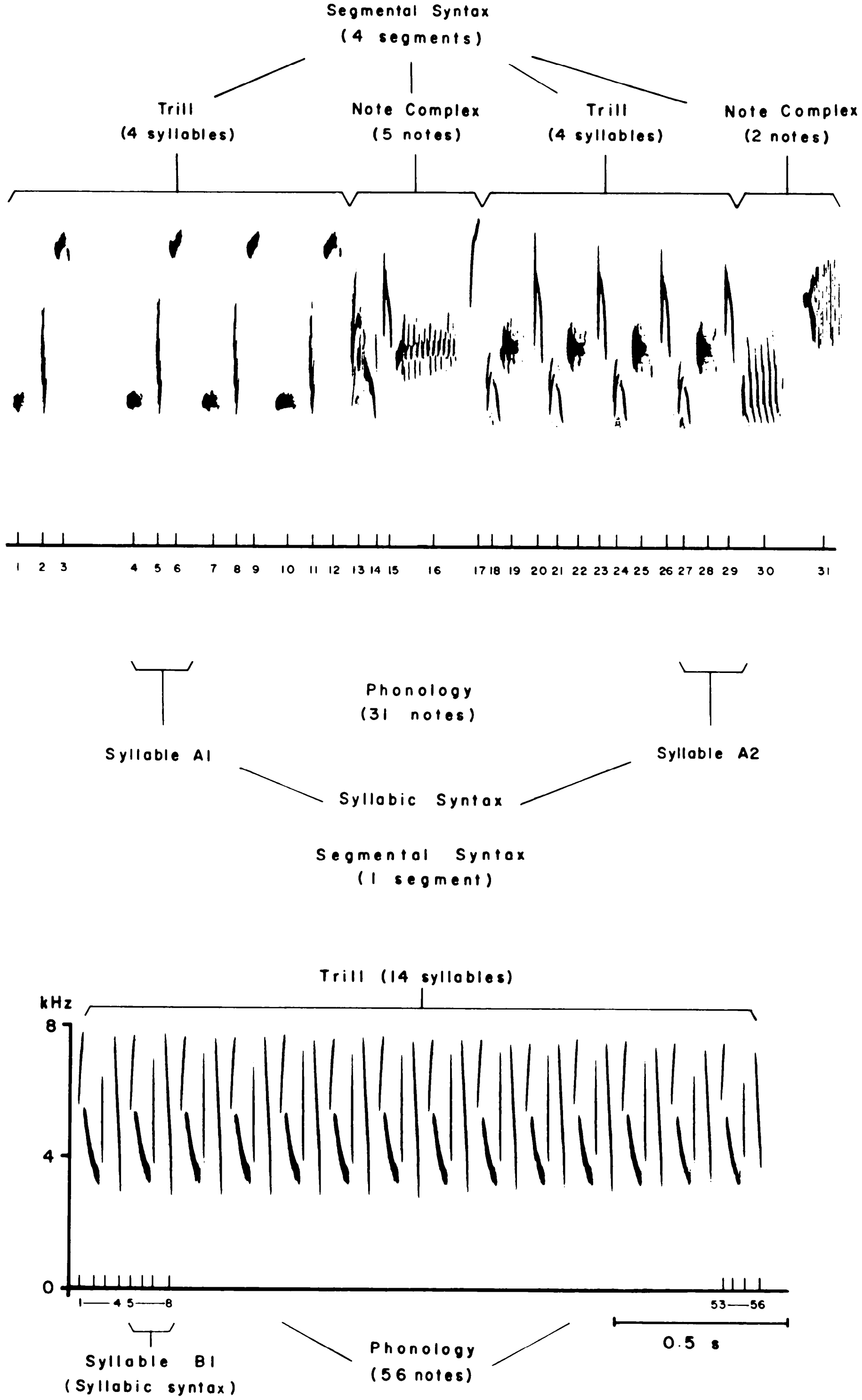
Acoustic Characteristics of Birdsongs

As a group, songbirds exhibit an astonishingly varied vocal repertoire. A comprehensive review of songbird acoustic phonetics is beyond the scope of this article. It is useful nonetheless to consider a fictional "typical" song as a means of providing some background into the kinds of sounds found in birdsong. To describe this song, several aspects of acoustic organization and structure must be addressed. On the prosodic level, a typical song might be 2 sec in duration, consisting of 50 or more discrete phonological units or "notes" (Figure 2). Notes can be roughly 10 to 100 msec in duration, and they are sometimes delivered very rapidly, often at rates up to five times faster than the rate of syllable production in normal human speech (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Marler, 1984). Groups of notes form "syllables" that are 100 to 200 msec in duration. Syllables can be grouped as phrases, and a song may comprise one phrase or many. In some songs, every note is different, but more often notes or syllables are repeated in regular patterns or "trills." The overall syntactical arrangement of phonological units is a species-specific trait. Field ornithologists often base their recognition of a species' song on the patterns of repetition within a song, without necessarily referring to the particular notes employed. This is the case with mockingbirds, thrashers, and catbirds, which incorporate into their song the sounds of other species, but deliver them with distinctive patterns of timing and repetition of their own (Boughey & Thompson, 1976; Baylis, 1982).

Our present discussion centers on the tonal quality or timbre of bird sounds. In this context, the acoustic structure of the notes themselves is the primary focus. Like a musical note, an utterance in birdsong may be described by (1) its duration, (2) its frequency, and (3) its amplitude. [Here we have chosen to use objective physical descriptors, such as frequency and amplitude, as opposed to their perceptual attributes, such as pitch and loudness, in order to avoid assumptions about how birds perceive sound. See Dooling (1982, in press) for reviews of what is known about how birds process and perceive acoustical information.] In addition, the description must encompass (4) the spectral organization of the sound (i.e., "timbre") and (5) the time-varying changes in the sound's properties over time (i.e., features of attack, decay, vibrato, etc., termed "sonance" by Seashore, 1938).

Compared to instrumental sounds or voice, the fundamental frequencies observed in birdsong are generally quite high, typically in the range of 3000 to 5000 Hz. A single song of a song sparrow, for example, might encompass a range of 2000 to 7000 Hz (Figure 2). Perhaps the greatest range within a single song is observed in the brown-headed cowbird, whose song begins with a low phrase at about 200 Hz and ends on a high note close to 11,000 Hz, a span of almost six octaves (Greenewalt, 1968; King & West, 1983).

A description of sounds as "bird-like" usually implies that they are both high pitched and also tonal, in contrast with the noisy, harsh quality that characterizes so many other animal sounds (Marler, 1977). Tonality is not a universal feature of bird sounds, however. In addition to songs, each species has a repertoire of 5–20 distinct calls, serving a range of functions from displays of hostility, the dissemination of alarm and the coordination


Segmental Syntax
(4 segments)
Trill
(4 syllables)
Note Complex
(5 notes)
Trill
(4 syllables)
Note Complex
(2 notes)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Phonology
(31 notes)
Syllable A1
Syllable A2
Syllabic Syntax
Segmental Syntax
(1 segment)
Trill (14 syllables)
kHz
8
4
0
1—4 5—8
53—56
0.5 s
Syllable B1
(Syllabic syntax)
Phonology
(56 notes)

of courtship and mating, to sounds that mediate in interactions between-parents and offspring (Nottebohm, 1975). Unlike songs, calls are typically brief, and many have a broadband, noisy structure. This common distinction between calls and songs can be clearly seen in a remarkable book by Bergmann and Helb (1982), which illustrates with more than 2000 sound spectrograms the songs and calls of some 400 species of European birds, about half of them songbirds. There is a predominant tendency for the oscine songs to be tonal. In most cases, not only are birdsongs tonal, but their constituent notes are in fact limited to a single, narrow-band "pure tone" (e.g., Figure 2), apparently devoid of the upper harmonics so characteristic of the voice and most musical instruments.

Even songs that sound dissonant or nasal to our ears may have a tonal basis. Members of the subfamily Icterinae range from the beautifully melodic songs of meadowlarks and orioles to the rasping buzz of the yellow-headed backbird. Sound spectrographic analysis reveals that this buzz consists, not of noise, but of a tone that is subjected to rapid and irregular frequency modulation (Figure 3). Because of variations in frequency of both the carrier and modulation, the resulting clusters of side-bands vary in interval, hence the noisy quality. The songs of different Icterid species can be arranged in a progressive series, ranging from the musicality of the Brewer's blackbird, through the pleasant dissonance of red-winged blackbird song, to the rather harsh buzz of the yellow-headed blackbird. Equivalent patterns of sidebands can readily be synthesized by different rates and depths of frequency modulation (Hund, 1942; Marler, 1969; Figure 4). By the appropriate choice of analysis bandwidth (see Figure 4), many apparently noisy birdsongs are found to be structured in this fashion (Marler, 1969).

Fig. 2. Sonagrams of typical song and swamp sparrow songs (Kay Digital Sona-Graph Model 7800, 8-kHz analysis range, 300-Hz analysis bandwidth). Like a musical score, sonagrams are "read" from left to right with distance on the horizontal axis corresponding to timing and duration. Position on the vertical axis indicates frequency, and amplitude is roughly indicated by the darkness of a trace at any point. The song sparrow song (above) comprises 31 separate notes (numbered). The notes are arranged as a *trill* of 4 syllables (each syllable made up of 3 distinct notes), followed by a *note complex* of 5 distinct notes, another *trill* of 4 3-note syllables, ending with a 2-note *note complex*. The segmental syntax of *trill–note complex–trill–note complex* is the most common pattern found in song sparrows. The swamp sparrow song (below) comprises 56 separate notes, arranged in a single *trill* of 14 syllables, each syllable composed of 4 distinct notes. The great majority of swamp sparrow songs are single trills, although some birds include a 2-trill song in their repertoire.

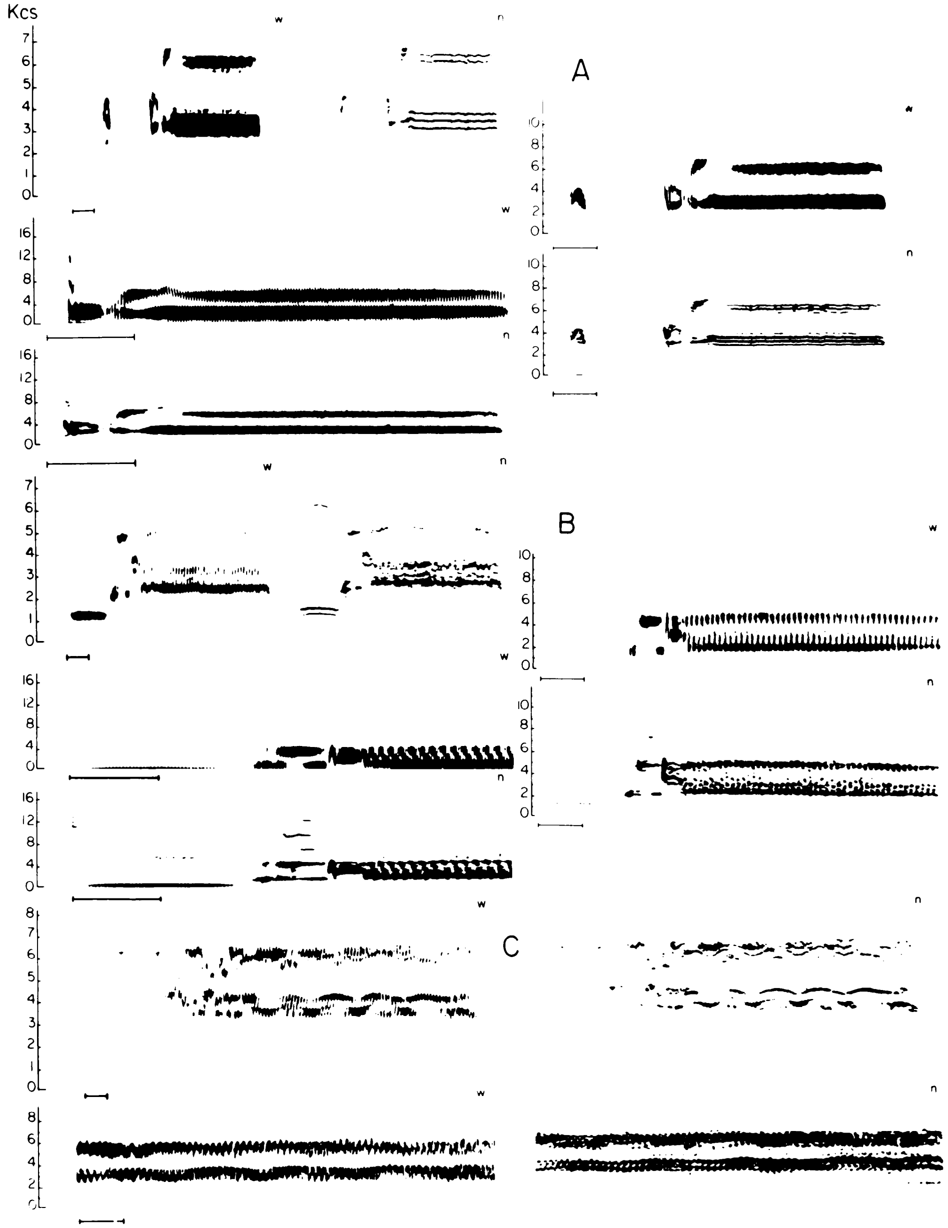


Fig. 3. Songs of three Icterids that include frequency modulated tones. The terminal portion of the Brewer's blackbird song (A) is pleasingly musical, while that of the red-winged blackbird (B) is nasal and dissonant, and that of the yellow-headed blackbird (C) is a harsh, jangling tone. A and B are shown at three playback speeds (all time markers shown are 0.1 sec); in each case, both wide filter bandwidth (w) and narrow filter bandwidth (n) analyses are shown, illustrating the effect of relative analysis bandwidth on the ability of the analyzer to resolve complex modulations (Kay Sona-Graph Model 6061B). Reprinted from Marler (1969, Plate 4).

Thus, whereas musical notes are generally steady-state, other than patterns of attack, decay, and vibrato, notes in birdsong are often subjected to extreme modulations across their duration. Modulations that are too rapid for our ears or analyzers to resolve result in broadband signals (Figures 3, 4), while slower modulations appear as impressive glissandi or crescendi (e.g., swamp sparrow in Figure 2). One of the more unusual acoustic features of birdsong is that amplitude modulations (AM) and frequency modulations (FM) are often coupled (Greenewalt, 1968). That is, the instantaneous amplitude and instantaneous frequency of notes covary. At lower frequencies, this coupling is direct such that amplitude increases as frequency increases. Above a certain frequency limit, which differs from species to species, the coupling remains equally as strong, but becomes inverse, with amplitude decreasing as frequency increases (Greenewalt,

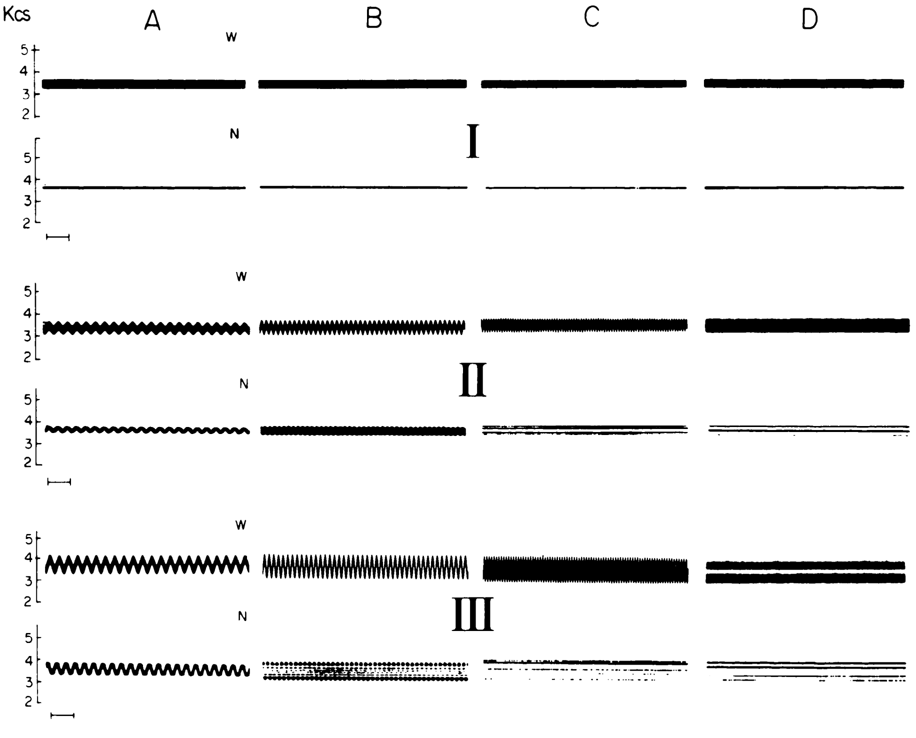


Fig. 4. A tone of about 3.5-kHz frequency modulated at various rates and depths and visualized sonagraphically (Kay Sona-graph Model 6061B). The modulation rates are: (A) 25 Hz; (B) 50 Hz; (C) 100 Hz; (D) 200 Hz. The modulation depth (or "range") in I is zero. In II and III, the depths are approximately ± 150 Hz and ± 250 Hz, respectively. In all cases, the analysis range is 8 kHz and both wide band (W, analysis bandwidth = 300 Hz) and narrow band (N, analysis bandwidth = 45 Hz) analyses of each sound are shown. The time marker is 0.1 sec. Modified from Marler (1969, Plate 3).

1968). This coupling between AM and FM accounts for many of the complex modulations observed in birdsong and provides a challenge for any mechanistic explanation of how song is produced.

In summary, the notes from which birdsongs are constructed vary widely in spectral structure, but are most typically pure tones with a narrow bandwidth, often subjected to varying degrees of FM and AM. At slower modulation rates, one observes pronounced glissandi and crescendi. With rapid modulation, either in frequency, amplitude, or both, there is an increase in bandwidth, resulting in complex tones and buzzes. In contrast with song, the absence of tonality from many bird calls implies that the vocal apparatus of songbirds can operate in several modes, at least one of which is specialized to generate sounds of a highly tonal nature.

Models of Songbird Phonation

The oscine syrinx (Figure 5) is located at the junction of the primary bronchi, beneath the bird's sternum, and just anterior to the heart. Originally called the "inferior larynx" (Hérissant, 1753; Cuvier, 1805) because of its position at the bottom of the trachea, the organ was later renamed the "syrinx" by Thomas Henry Huxley (1871). Despite early attention to syringeal structure and function, relatively little contemporary work was done on phonation in birds until interest in the problem was renewed with publication of the book *Bird Song: Acoustics and Physiology* by Crawford Greenewalt in 1968. A review of the anatomy and the experimental information available from earlier studies led Greenewalt to argue persuasively that an accurate theory of syringeal operation could be deduced largely from acoustic analyses of bird sounds. The elegant model he proposed in this way has endured with only minor modifications as the most widely held view of how birds sing (Stein, 1968; Gaunt & Wells, 1973; Brackenbury, 1980, 1982; Gaunt & Gaunt, 1985).

Greenewalt's (1968) model embodies three major features: (1) A pair of membranes, the medial tympaniform membrances (MTMs), act as acoustic sources when induced into vibration by airflow; (2) the MTMs, located separately on the medial walls of the two bronchi, are viewed as functionally independent—that is, songbirds are thought to have two "voices" capable of simultaneously producing two unrelated tones; and (3) all modulations occurring in birdsong are thought to be source-generated, resulting from changes in the state of the MTMs. Greenewalt (1968) specifically argued against the possibility of modulation due to vocal tract resonances that would modify the output of the syringeal sources.

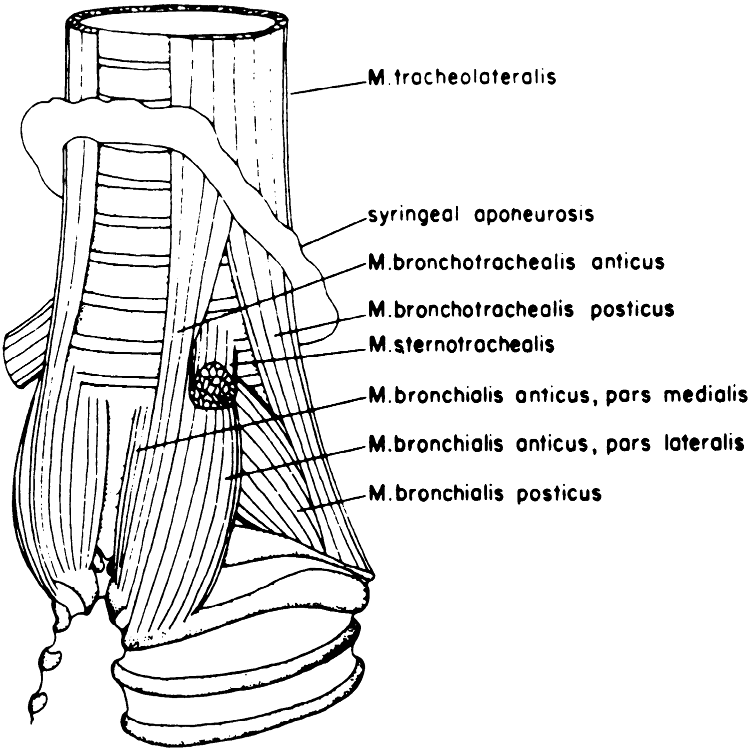


Fig. 5. External view of a typical oscine syrinx, showing the arrangement of cartilaginous elements and syringeal musculature. Opening at the top is the base of the trachea and the incomplete rings at the bottom are the first several elements of the two bronchi leading to the lungs. Each of the seven syringeal muscle pairs is labeled, as is the syringeal aponeurosis, the point at which the trachea emerges from the membrane of the interclavicular airsac in the thoracic cavity. Modified from George and Berger (1966, Figure IX.19).

These three points, and the evidence adduced in their support, are reviewed below. More comprehensive discussions of this material can be found in reviews by Brackenbury (1980, 1982) and Gaunt (Gaunt & Gaunt, 1985; Gaunt, 1987).

THE ACOUSTIC SOURCES

The songbird's syringeal skeleton is characterized by fusion of the lowest of the cartilaginous rings that encircle the windpipe or trachea. The pessulus, a saddlelike extension of the lowest fused ring, projects down and bisects the tracheal lumen to form two openings where the two bronchi from the lungs attach to the syrinx (Figure 5). Several cartilaginous rings below this junction are incomplete, or C-shaped, with the open portion of the C facing medially (Ames, 1971; Warner, 1972). Of the several membranes and lobes of tissue within the syrinx that could play a role in the induction of vibration (Figure 6), most attention has focused on the medial tympaniform membranes (MTMs). These are very thin, elastic membranes stretching across the medial surfaces of the bronchi, connected on their ventral and dorsal aspects to the ends of the bronchial C-rings. Opposite the MTMs are the external labia, thickened pads of epithelium on the inner lateral walls of the third bronchial rings (Häcker, 1900; Setterwall, 1901; Miskimen, 1951; Greenewalt, 1968; Warner, 1972).

The syrinx is enclosed entirely within the interclavicular airsac (ICAS), one of the several airsacs that branch from the bird's lungs and ramify throughout its body (Ames, 1971; Schmidt-Nielsen, 1971; Kloek & Casler, 1972). Songbirds generally are unable to sing if the ICAS has been punctured (Hérissant, 1753; Smith, 1977), suggesting that pressure must be maintained in this air space for normal sound production to occur. Rüppell (1933) confirmed this idea by isolating a freshly dissected herring gull syrinx in a bell jar, with tubes attached to permit airflow through the inside of the organ. He observed that the MTMs only vibrated when extended into the lumen of the bronchi as a result of increased external air pressure such as is presumably generated within the interclavicular airsac.

Greenewalt (1968) suggested that, when the MTM extends into the bronchus, airflow from the lungs creates a Bernoulli effect, lessening pressure on the bronchial side of the membrane and drawing it in even further. As the membrane extends maximally into the bronchus, the Bernoulli force reaches an equilibrium with the counteracting force of the membrane's own elastic tension. At this point, viscous flow losses reduce Bernoulli's force and the membrane withdraws until forces reequilibrate and the cycle repeats. Thus, Greenewalt proposed that oscillation of the syringeal membranes is aerodynamically driven, similar to the way oscillation is induced in human vocal folds (Ladefoged, 1962; Lieberman, 1977, 1984).

There are two critical differences between the syringeal membranes of the bird and vocal folds of the human larynx as acoustic sources. First, the two vocal folds act as a single acoustic source, whereas the MTMs are thought to be independent of each other. Second, the oscillatory cycle of the vocal folds includes a significant period during which the two folds are in contact, occluding the airstream completely. Closure of the vocal folds is responsible for creating discrete "puffs" of air in the glottal flow pattern, accounting for the harmonic nature of the human vocal source (Fant, 1960; Broad, 1973). By contrast, it was suggested that the MTMs rarely traverse the entire bronchial lumen and thus create a more nearly sinusoidal pressure disturbance. The significance of a sinusoidal acoustic source is that it would account for the pure tone nature of many bird sounds (Greenewalt, 1968).

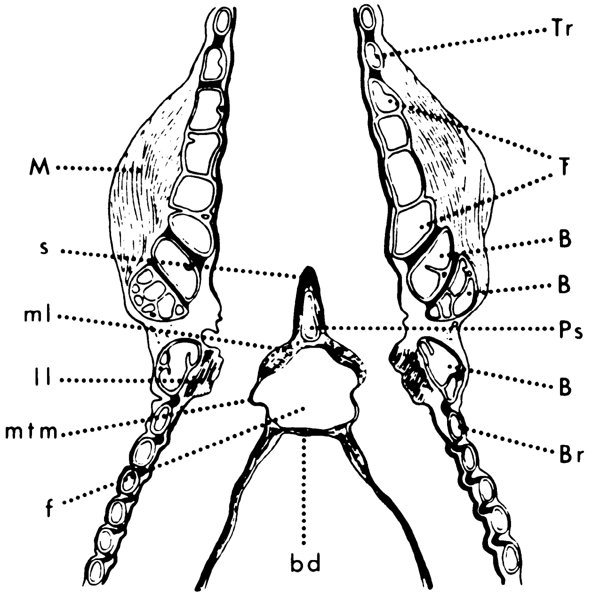


Fig. 6. Cross section through a typical oscine syrinx, in this case a European blackbird. Abbreviations: Tr, tracheal ring; T, fused tracheal rings; B, modified bronchial ring; Br, bronchial half-ring; Ps, pessulus; M, musculature; s, semilunar membrane; ml, medial labium; ll, lateral labium; mtm, medial tympaniform membrane; f, foramen continguous with interclavicular airsac; bd, bronchidesmus (interbronchial ligament). Modified from Häcker (1900, Figure 2).

This model of how birdsong is produced has little in common with the acoustic sources found in wind instruments, most of which are described as pressure- or flow-controlled valves (Benade, 1976). A recent alternative to the Greenewalt model, however, suggests that birds generate sound in a manner similar to a hole-tone whistle (Nottebohm, 1976; Gaunt, Gaunt, & Casey, 1982; Casey & Gaunt, 1985). This hypothesis was proposed in part because theoretical analyses of membrane vibration did not support Greenewalt's (1968) view that the MTMs would oscillate in an essentially sinusoidal fashion (Casey & Gaunt, 1985; Gaunt & Gaunt, 1985). A hole-tone whistle, also known as a "Rayleigh bird call" (Rayleigh, 1896, v. II, p. 410), generates sound by creating stable vortices in the air flow that are shed as the result of a constriction. Such a whistle is capable of creating a relatively pure tone (Rayleigh, 1896; Chanaud & Powell, 1965; Wilson, Beavers, DeCoster, Holger, & Regenfuss, 1971; Casey & Gaunt, 1985).

THE TWO-VOICE THEORY

The most striking anatomical feature of the songbird syrinx is that almost all components are present in duplicate. Not only does each bronchus have an identical set of membranes and muscles, but each side is also separately innervated by the descending branches of the left and right XIIth cranial (hypoglossal) nerves. The development of the sound spectrograph as a technique for visualizing acoustic patterns initially led to the theory that birds can produce two independent sounds simultaneously (Potter et al., 1947; Borror & Reese, 1956). An outstanding example is seen in the song of the wood thrush (Figure 7), characterized by Greenewalt (1968) as the most accomplished and versatile "internal duettist." Its loud, melodious song is described as liquid, clear and flutelike (Saunders, 1935). In the terminal trill of a wood thrush song, not only are two nonharmonic tones clearly observed, but these two "voices" often appear to be modulated in different ways. Through the use of sonagrams, and frequency filters to isolate apparently separate "voices" in birdsong, Greenwalt (1968) demonstrated that the "two-voice" phenomenon is widespread among different families of birds.

The acoustic evidence for functional independence of the two syringeal sources was supported later by the physiological experimentation of Nottebohm (e.g., 1971, 1972), who sectioned the nerve bringing motor impulses to one or the other syringeal side, resulting in a predictable distortion or loss of certain acoustic elements in a complex song. For example, when the nerve leading to the left side of a canary's syrinx was sectioned, Nottebohm and Nottebohm (1976) discovered that a majority of syllables dropped out of its postoperative songs. Sectioning the right side alone had relatively little effect on most of the syllables, but caused a few syllables (presumably those that were unaffected by a left nerve section) to drop out. In each case the elements deleted by right- or left-side section are presumed to have been generated by the disabled side of the organ. Similar results have been obtained from a variety of songbird species (Nottebohm, 1971; Lemon, 1973; Nottebohm & Nottebohm, 1976; Seller, 1979). On the basis of these findings, it became generally accepted that songbirds have independent control of their two syringeal sources. Intriguingly, the left side has most often proved to be dominant (Nottebohm, 1977, 1980; Arnold & Bottjer, 1985).

MODULATIONS IN BIRDSONG

Divergent examples of how acoustic resonances might affect the output of a sound source are provided by musical wind instruments and the human voice. The major distinction to be drawn here is as follows: In the musical

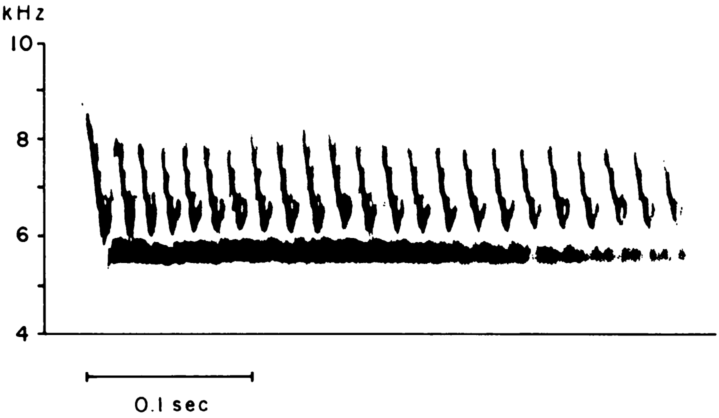


Fig. 7. Sonagram of the terminal trill of a wood thrush song (Kay Digital Sona-Graph Model 7800, 16-kHz analysis range, 300-Hz analysis bandwidth). It illustrates the simultaneous presence of two "voices" that are harmonically unrelated and independently modulated, but note that they are pitched within the pass band of a single vocal tract resonance.

instrument example, the source and acoustic resonator are strongly coupled such that the source is constrained to vibrate only at frequencies permitted by the associated resonances. In human speech, by contrast, the source (the vocal folds of the larynx) and the associated resonances of the vocal cavities act independently such that the source is not directly influenced. Instead the resonances act as a complex acoustic filter, selectively emphasizing particular frequency bands and attenuating others. In both cases, the emitted sound is modulated by changes in acoustic resonances.

The fact that the two sides of an oscine bird's syrinx appear to operate independently originally led to the hypothesis that vocal tract resonances play no role in the production of birdsong (Greenewalt, 1968). The avian vocal tract could not be acting as a coupled resonator in the manner of a musical instrument, since a single resonator, the vocal tract, could not support two harmonically unrelated sounds. It was also argued that vocal resonances could not be acting as an uncoupled acoustic filter, in the manner of human speech, because no relationship was found between the amplitude and frequency of song syllables and the predicted resonances of the trachea of songbirds modeled as a simple tube closed at one end. Similarly, analysis of the broadband glissandi found in many birdsongs uncovered no relationship between amplitude and frequency corresponding to these inferred tube resonances (Greenewalt, 1968). This argument against acoustic filtration by vocal resonances assumes that the filter properties of a bird's vocal tract are static, in contrast to humans where resonances are actively and rapidly changed during phonation. It becomes less compelling if we admit the possibility of equally dynamic changes in the configuration of the vocal tract of the bird.

The implication that vocal tract resonances have no influence at all over the acoustic features of birdsong was puzzling, since any tube or cavity with relatively small acoustical losses has resonant properties. To explain this apparent paradox, Greenewalt enlisted the collaboration of J. L. Flanagan of Bell Laboratories. They proposed a theoretical model demonstrating that, for a certain ratio of tube diameters between the syringeal constriction at the vibration source and the lumen of the trachea, the impedances of the source and the trachea will be matched, creating an acoustical loss at this boundary. In this special case, one expects no tracheal resonance reinforcement or attenuation. One species, the song sparrow, was found to conform to the predicted measurements (Greenewalt, 1968).

Another reason for not invoking resonance modulation in birdsong was the ability of the model of the syringeal sources proposed by Greenewalt to account for all the modulations observed in song. All the variations observed could be viewed as source-generated. Greenewalt's model also provided an explanation for the unique coupling observed between AM and FM, as described above. It was argued that the intricate musculature surrounding the syrinx, consisting of seven separate pairs of muscles, with an identical set on each side, is so situated as to exert complete control over the relative tension of the MTMs. As with any membrane, the frequency of vibration is determined by tension. Thus, the frequency modulations that occur in birdsong, superimposed on the pure tone that serves as a carrier, could be explained simply by invoking complete, if subtle, control over membrane tension by this musculature.

The coupling between FM and AM was explained as follows: With weak tension on the MTM, the membrane vibrates at low frequencies, but it also extends maximally into the bronchial lumen. The amplitude of vibration is thus limited by proximity to the opposite wall. As tension increases, the frequency increases, and so does the amplitude of vibration, since there is now a greater average physical distance between the MTM and the opposite bronchial wall, allowing greater displacement during oscillation. Frequency and amplitude are thus directly coupled at relatively low frequencies. As tension on the MTM increases, a point will be reached where the membrane oscillation no longer traverses a significant distance across the lumen on the bronchus. Direct current (DC) "slippage" of the airstream will occur and the acoustic amplitude will decrease. As the frequency increases further, the amplitude decreases. Thus, at relatively high frequencies, FM and AM are inversely coupled, whereas at low frequencies they are postively coupled.

New Perspectives on Syringeal Function

Most contemporary investigators agree that the details of syringeal mechanics, especially the origin and maintenance of vibration in the syringeal sources, are still only poorly understood (Gaunt & Gaunt, 1985; Gaunt, 1987). Greenewalt (1968, p. 184) himself suggested that "direct experimental evidence is much to be desired." Meanwhile, the three major premises of Greenewalt's original model—(1) that the MTMs are the acoustic sources, (2) that the two MTMs are independent, and (3) that all modulations are source-generated—have generally been regarded as fundamental features of syringeal operation which must be incorporated into any interpretation of syringeal function.

The choice was made to begin experimentation by reexamining these fundamental assumptions. The results suggest that there is a need for revision or even the complete reversal of earlier views. To date, the results address especially the two-voice theory and the role of vocal tract resonance in phonation. Although the nature of the acoustic sources has not yet been directly explored, the experimental results on other issues suggest that a wide range of theoretical issues must be reconsidered.

EMENDATIONS OF THE TWO-VOICE THEORY

The "dee" note of the "chick-a-dee" call of the black-capped chickadee is not a pure tonal sound, but it is instead broadband. Its spectrum appears as a series of overtones, evenly spaced at roughly 400-Hz intervals (Figure 8), suggestive of a harmonic series, but with its fundamental and first several harmonics missing (Greenewalt, 1968). The lowest frequency components with significant energy consistently occur at about 1600 and 2000 Hz (Components #1 and #2 in Figure 8). One way to investigate the contributions of the two "voices"to such a sound is to cut the nerve leading to one side of the syrinx or the other (cf., Nottebohm & Nottebohm, 1976). If the two sides of the bird's syrinx operate independently in the production of this sound, then one expects that left or right unilateral syringeal denervation will lead to a predictable loss of all or some of these frequency components; all or none should be lost if one side alone is normally responsible for generating the signal, and only some if the two sides each contribute a subset of the components, the sum of the two contributions constituting the complete song. In either case, a linear summation of signals obtained from left and right operated birds should result in a close approximation of the normal signal, as appeared to be the case in earlier work (Nottebohm & Nottebohm, 1976).

This predicted result was not obtained. Instead, unilateral denervation in the chickadee leads to a noisy sound in which is embedded a strong harmonic signal. Depending on whether the right or left nerve is sectioned, the fundamental frequency of the postoperative signals corresponds either to the frequency of the first or second spectral component of the normal signal (Figure 9). The sum of harmonics of these two signals does not yield the spectral profile of the intact signal (Nowicki & Capranica, 1986a). That is, the linear addition of left and right postoperative signals does not reproduce the regular pattern of the normal signal. The implication is that neither side of the syrinx acts alone in production of the "dee" sound. Nor do the two sides act together but with each side contributing an independent subset of the signal. Thus, the chickadee's two syringeal sources must be coupled in some fashion (Nowicki & Capranica, 1986b).

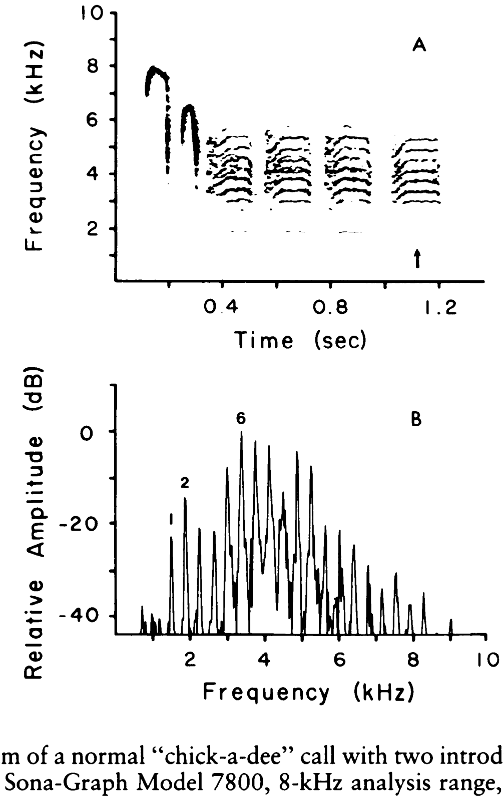


Fig. 8. (A) Sonagram of a normal "chick-a-dee" call with two introductory and four "dee" notes (Kay Digital Sona-Graph Model 7800, 8-kHz analysis range, 45-Hz analysis bandwidth). (B) Amplitude spectrum for a 40-msec section of a "dee," as marked with an arrow in A (Nicolet Mini-Ubiquitous FFT analyzer; frequency resolution, 25 Hz). The first two components with significant energy and the component of maximum amplitude are numbered, as described in the text. Reprinted with permission from Nowicki and Capranica (1986a, Figure 1). Copyright 1986 by the AAAS.

These results present an interesting puzzle. The harmonics observed in postoperative calls presumably arise from the action of the acoustic source on the intact side (contralateral to the nerve section). But, a summation of left and right postoperative signals yields fewer frequency components than observed in the normal signal. The problem, then, is to account for the production by an intact bird of frequencies that are absent when either syringeal source is acting alone.

One clue is found in the correspondence between the postoperative fundamentals and the first two spectral components of the normal signal (Figure 9). The fact that the frequency difference between these components invariably equals the frequency spacing between higher overtones of the normal signal suggests that the orderly spectral components of the normal call may arise as heterodyne frequencies resulting from a nonlinear interaction between the two syringeal sources. The simplest such interaction is amplitude modulation. Like all nonlinear systems, AM between two signals yields not only the original frequencies of the modulated signals, but also new frequencies that are derived as sums and differences of the original frequencies (Nowicki & Capranica, 1986b). Such heterodyne frequencies, also generated by frequency modulation (cf. Marler, 1969, p. 5), are known as sidebands.

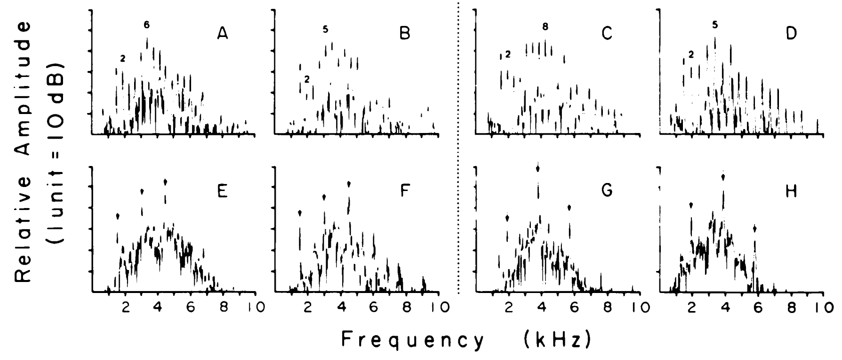


Fig. 9. Comparison of pre- and postoperative amplitude spectra for four different chickadees. (A–D) Preoperative spectra, with frequency components numbered as in Figure 8B. (E,F) Postoperative spectra following section of the nerve leading to the right side of the syrinx. (G, H) Postoperative spectra following section of the left nerve. Arrows indicate postoperative harmonic components as described in the text. Each postoperative signal was generated by the same bird responsible for the preoperative signal above it. Note the relationship between preoperative Component #1 and the postoperative fundamental in the case of a right nerve section and between preoperative Component #2 and the postoperative fundamental in the case of a left section. Reprinted with permission from Nowicki and Capranica (1986a, Figure 2). Copyright 1986 by the AAAS.

A mathematical model of amplitude modulation between two sources producing harmonic signals supports the hypothesis that such an interaction is responsible for the generation of the "dee" sound (Nowicki & Capranica, 1986b). By substituting observed values derived from postoperative signals, with fundamental frequencies of 1600 and 2000 Hz, a series of heterodyne frequencies is generated that corresponds precisely to the frequency spectrum observed in a normal "dee" note.

The physical basis of this source–source interaction remains conjectural. It is likely to entail a passive physical principle, such as pressure coupling between the two MTMs via the interclavicular airsac, which is regulated by bilateral neural control. Also, it is evident that the chickadee is capable of turning this bilateral syringeal interaction "on" and "off," since other notes in the "chick-a-dee" call, as well as different calls in the species' repertoire, lack sidebands and are more typically pure tonal.

Given that the acoustic structure of the "dee" is unusual, to what extent might the bilateral syringeal interaction observed in its production be found in the production of pure tone sounds by this or other species? Preliminary results suggest that sparrows might also employ some form of heterodyning in the production of broadband signals (Nowicki & Capranica, 1986b; S. Nowicki, P. Marler & S. Peters, unpublished observations), and that the production of narrow-band or pure-tonal sounds might involve the active suppression of bilateral interaction (S. Nowicki, unpublished observations). It begins to look as though the capacity for interactions between the two syringeal sides may be a ubiquitous feature of song production, responsible for some of the more unusual modulations and spectral properties of oscine birdsongs.

Other investigators have recently questioned the universality of the two-voice theory on independent grounds. An alternative approach to disabling one side of the syrinx is to block airflow to that side by physically plugging the bronchial passage below it. McCasland (1987) performed this procedure on canaries, a species for which earlier data had suggested that the two syringeal "voices" operated in an independent and lateralized fashion (Nottebohm & Nottebohm, 1976). Instead of finding that certain elements dropped out or were distorted, as would be predicted from the two-voice theory, he found that all syllables in a song were distorted to some extent, regardless of which side was plugged (McCasland, 1987). As with denervation experiments, the physiological interpretation of bronchus plugging is compromised by the paucity of information about how acoustic oscillation is induced in the syrinx. From these experiments and our own, however, we are more and more inclined to conclude that previous statements of the two-voiced theory encompass only a fraction of the true versatility of the syrinx with regards to the possible action of its acoustic sources.

A REAPPRAISAL OF TRACHEAL RESONANCE IN SONG PRODUCTION

Acoustic resonances are determined by the size, shape, elasticity, and other physical properties of the resonant cavity and especially by sound velocity. Thus, the density of the propagating medium has a direct effect on resonance frequencies. When atmospheric nitrogen, comprising about 80% of normal air, is replaced with the less dense gas helium, the velocity of sound increases from 331 to 578 m/sec. Accordingly, one predicts a 74% increase in the resonance frequencies of a simple tube (Hardy, Telfair, & Pielemeier, 1942; Beranek, 1949). The effect of this change on a sound-producing system will vary qualitatively, however, depending on the manner in which resonances interact with the acoustic source, as already illustrated by the divergent examples of wind instruments and speech.

When a human speaks in a helium atmosphere, the fundamental frequency of vocal fold vibration *does not* change appreciably (Beil, 1962; Holywell & Harvey, 1964). Instead, the perceived increase in pitch, making a person sound like Walt Disney's Mickey Mouse, is due to an upward shifting of the vocal tract formants with a resulting emphasis on higher frequency overtones. The alternative case is illustrated when a wind instrument, for example a trombone or trumpet, is played in helium. Here the fundamental frequency of the tone played *does* increase, in the same proportion as the shift calculated for the resonances of the instrument's bore. The brass player's lips are strongly acoustically coupled to the horn's bore, and the mode of vibration of the source is determined by the tube's resonances. The same is true of the vibrating reed in woodwinds (Helmholtz, 1877; Benade, 1976).

If acoustic resonances play no role at all in bird phonation, as is assumed by the Greenewalt model, then song should remain unaffected when a bird sings in a helium atmosphere. This prediction was not confirmed by experiment. The songs and calls of nine species of oscine birds were recorded both in normal air and in a helium atmosphere (Nowicki, 1987). The results were equivalent for all species and will be illustrated here by the song sparrow, the swamp sparrow, and the black-capped chickadee. The most striking effect of helium is that a harmonic overtone is added to each song element appearing as a pure tone in normal atmosphere (Figure 10). These songs also show a marked overall decrease in amplitude, often in excess of 10 decibels (dB). Sounds with a broad band of frequencies, such as the chickadee's "dee," display a shift in their amplitude spectra, with higher frequencies emphasized (Figure 11).

The occurrence of any change at all in the helium atmosphere proves that resonances are involved in determining the acoustic properties of birdsong, reversing the long-standing conclusion to the contrary. But what role do these resonances play? In no case was an appreciable shift in the fundamental frequencies of a song observed, ruling out the possibility that vocal resonances control the vibration characteristics of the syringeal source in the manner of a musical wind instrument. The shifted emphasis to higher frequencies in the chickadee's note (Figure 11) is analogous to the change ob-

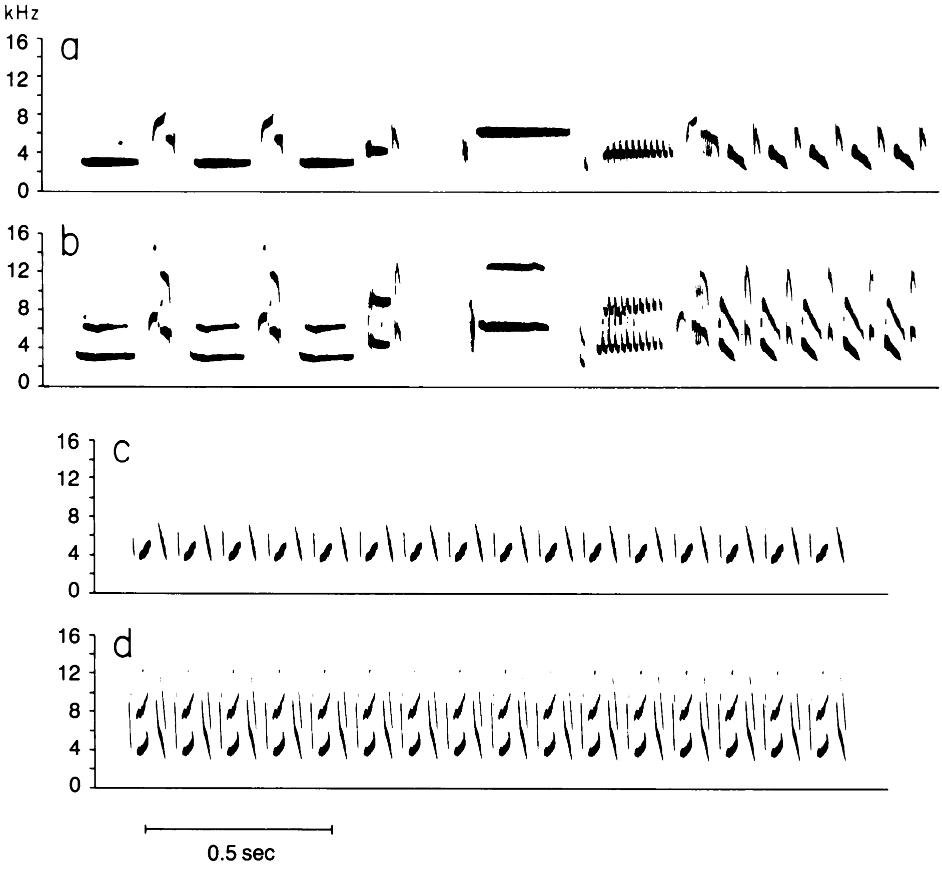


Fig. 10. Sonagrams of typical song and swamp sparrow songs recorded both in normal air and in helium air. (a) Song sparrow song, normal air; (b) same song sparrow songtype, but in helium air; (c) swamp sparrow song, normal air; (d) same swamp sparrow songtype, but in helium air (Kay Digital Sona-graph Model 7800, 16-kHz analysis range, 300-Hz analysis bandwidth). Note that every note appearing as a pure tone in normal air displays harmonic overtones in helium.

served in the human voice, with a filter function shifting in the direction and magnitude expected of changed resonances in the vocal tract. Other species' broadband sounds, whether tonal or noisy, behave similarly in helium (S. Nowicki, unpubl. data).

The pure tone examples (e.g., Figure 10) are of greatest interest, since they represent the most uniquely "bird-like" sounds. Here, too, the changes observed are comprehensible if the bird's vocal tract is acting as an acoustic band-pass filter that is not significantly coupled to the vibration source, as in the human voice. The argument, illustrated with a song sparrow song that includes a series of unmodulated pure tone notes (Figures 12, 13), suggests that the "whistles" so commonly observed in birdsong are not generated as pure tones at all. Rather, they are harmonic sounds with relatively high frequency fundamentals. Normally, the energy in their second harmonics is strongly attenuated by a filter centered around the fundamental frequency. In helium, this filter is shifted upward, revealing the harmonic content (Nowicki, 1987; Figure 14). Since neither the fundamental frequency nor the second harmonic is now squarely centered on the best frequency of the vocal filter, both have reduced amplitudes compared to the fundamental alone in ordinary atmosphere.

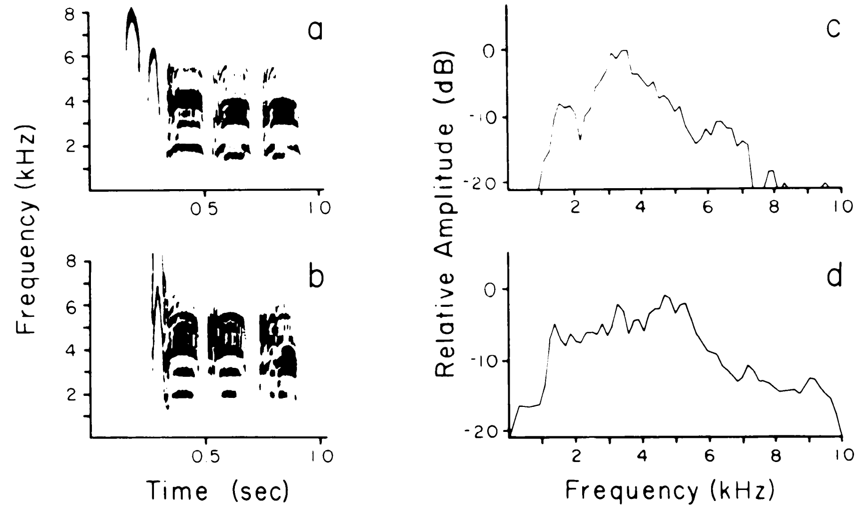


Fig. 11. (a, b) Sonagrams of "chick-a-dee" calls in normal air (a) and helium air (b) (Kay Digital Sona-Graph Model 7800, 8-kHz analysis range, 300-Hz analysis bandwidth). (c, d) Amplitude spectra of the second "dee" notes of the calls illustrated in a and b, respectively (spectra generated digitally, using a 25-kHz sampling rate, 10-kHz analysis range, 128 point FFT, yielding a frequency resolution of 156 Hz; each spectrum is an average calculated over the entire duration of a note). Note the shifted emphasis to higher frequencies in the helium case (b, d).

Evidence that the vocal tract filter is not simply passive, but is actively coordinated with syringeal output, comes from cases in which the second harmonic of a syllable falls in the same frequency range as the fundamental of an adjacent syllable (Figures 12, 15). The harmonic would appear in normal air if the frequency response of the filter were invariant from syllable to syllable. Filtering out this harmonic, which clearly occurs, can only be accomplished if the center frequency of the filter shifts rapidly prior to production of that syllable.

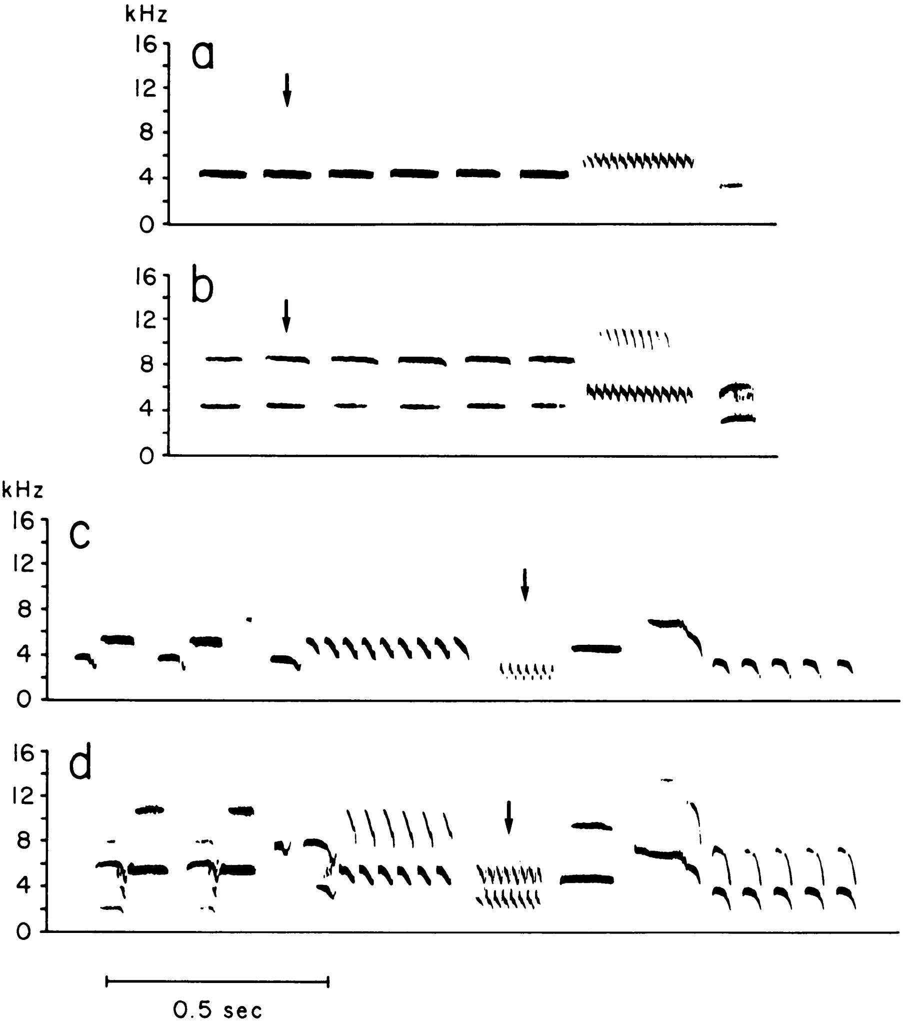


Fig. 12. Sonagraphs of two song sparrow songtypes recorded in normal air (a, c) and helium air (b, d) (Kay Digital Sona-Graph Model 7800, 16-kHz analysis range, 300-Hz analysis bandwidth). Arrows mark notes illustrated in Figures 13 and 15, as discussed in the text.

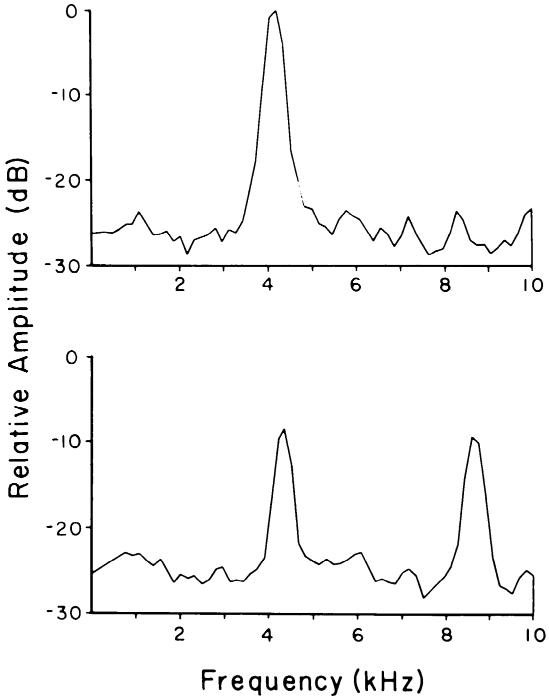


Fig. 13. Amplitude spectra of notes marked with arrows in Figure 12a, b (above is normal air, below is helium air), as described in the text. Note the lack of a significant shift in the fundamental frequency as well as the appearance of a true harmonic in helium. The amplitudes of signals in helium have been increased in gain by 10 dB compared to their counterparts in normal air to best illustrate the emergent frequency component. Spectra are averages generated digitally as described for Figure 11c, d. Modified from Nowicki (1987, Figure 2).

In theory, a bird can adjust its vocal filter in several ways. The resonance frequences of a tube, such as the trachea, vary not only with its length, but also with the degree to which the open end is constricted or flared out (Benade, 1976). Constriction of the bird's laryngeal opening (the glottis) could effectively lower the resonance frequency of the vocal tract. Similarly, the degree to which the vocal tract is flared, as by opening the beak, will influence the resonance frequencies. Preliminary data support the idea that postural changes in a bird's beak and throat position reflect changes in "tuning" of a vocal filter. Swamp sparrows were videotaped in close-up profile during song, with an oscillogram of the bird's song simultaneously displayed. An analysis of postural angles suggests that beak opening correlates with frequency, not amplitude, of the sound produced (S. Nowicki,

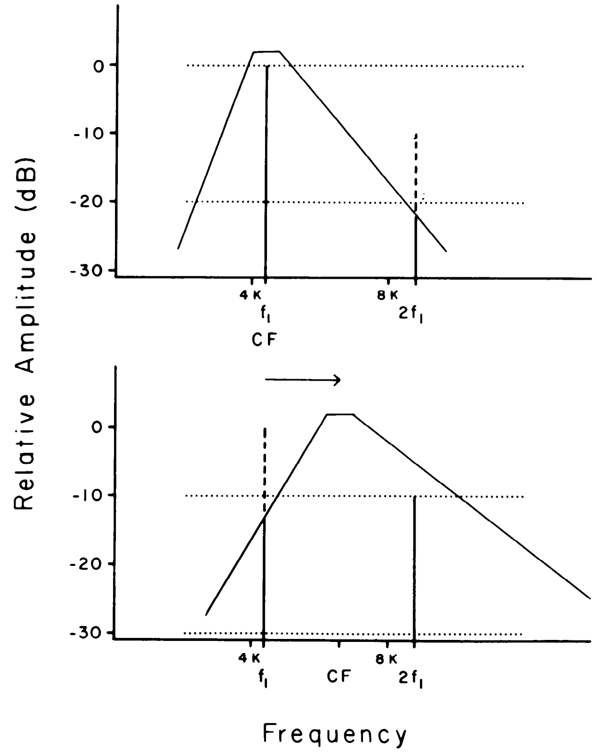


Fig. 14. Schematic representation of the predicted effect of helium on a hypothetical vocal tract filter. In normal air (above), the filter is centered over the fundamental frequency, effectively attenuating the second harmonic component. The filter's center frequency is shifted upward by 67% in the helium atmosphere (below), attenuating the fundamental component and revealing the second harmonic (Nowicki, 1987). The frequency components approximate those illustrated in Figure 13. The filter roll-off approximates a 4 pole Butterworth filter (24 dB/octave). The assymmetry is the result of representing the filter function on a linear frequency axis to correspond with Figure 13. The horizontal dotted lines approximate the dynamic range of a typical sonagraph, demonstrating how the residual energy in the second harmonic in normal air might not be observed in a sonagraphic analysis.

unpubl. data). That is, the bird does not appear to open its beak wider for higher amplitude sounds, but does open it wider for higher frequency sounds, as predicted if the beak is a partial determinant of acoustic resonances of the tract.

These results demonstrate that by no means are all of the acoustic properties of birdsongs source generated. Like human speech, the vocal tract is actively involved in shaping the emitted sounds. There are significant contrasts as well, perhaps the greatest being that oscine birds possess two sound

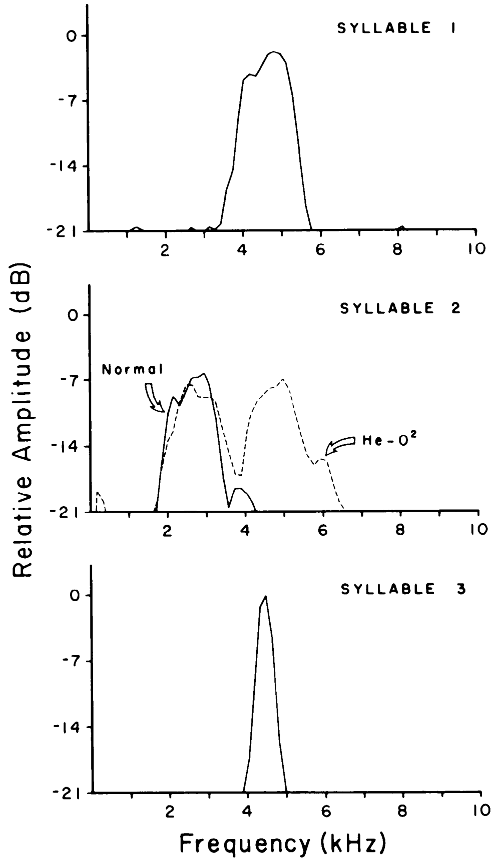


Fig. 15. Amplitude spectra of three consecutive syllables in normal air (solid lines), illustrating the overlap in frequency ranges of the second harmonic component of Syllable 2, apparent only in helium (dashed line), and the normal fundamental frequencies of adjacent Syllables 1 and 3. Syllable 2 here corresponds to the syllable marked with arrows in Figure 12c, d. The amplitude spectrum in helium is shown for Syllable 2 only. Note that a change in the frequency response characteristics of the vocal filter must occur from syllable to syllable in order to account for the lack of energy in the second harmonic component of Syllable 2 in normal air (spectra are averages generated digitally as described for Figure 11c, d).

sources, as opposed to the single source of the of the human larynx. If these two sources independently produce sounds, as appears to be true in at least some cases (e.g., Figure 7), then both must fall within the band-pass frequency range of the vocal filter. Further, the syringeal sources produce sounds with relatively high fundamental frequencies, on the order of several thousand Hertz, as opposed to fundamentals of less than 200 Hz characteristic of speech. Thus, the bird's vocal filter tends to concentrate energy at a single frequency during the production of tonal songs, whereas the human vocal tract varies the distribution of energy in a more complex fashion over a wide range of frequencies. When a bird produces calls rather than songs, however, it evidently often moves into a mode of operation more like that of the human voice.

It may be that a closer analogy for the role of a vocal filter in birdsong is found not in human speech, but instead in human singing. In the human soprano voice there is considerable overlap between the fundamental frequency of the source and the lowest resonance (or formant) of the tract (Sundberg, 1977). Such overlap occurs normally in birdsong (Nowicki, 1987). Whereas the laryngeal source and the vocal tract filter are generally thought of as acting completely independently during normal speech (Fant, 1960), it is now recognized that because of this overlap, interactions between the source and the vocal tract are important determinants of voice quality in the singing voice (Rothenberg, 1981, 1987a, b). These interactions reinforce the acoustical power of the fundamental frequency, leading to the efficient production of an even purer tone than would be possible with a simple acoustic filter, reminiscent of the narrow-band frequencies so often observed in birdsong. As we learn more about mechanisms of human singing, it may be that more parallels with birdsong will emerge.

NEW REQUIREMENTS FOR A MODEL OF THE SYRINGEAL ACOUSTIC SOURCES

A complete model of how acoustic oscillation arises in the syrinx, whether through the action of the MTMs or other anatomical elements, must await the implementation of techniques for directly monitoring vibrations in an intact, singing bird. Meanwhile, three new points can be made about how the syringeal sources appear to operate.

First, the sources themselves do not necessarily have to be capable of generating pure tones, as was previously thought. One major criticism of Greenewalt's (1968) model of MTM vibration is the unlikelihood that a vibrating membrane could produce a pure tone, as opposed to a more complex signal (Gaunt et al., 1982; Casey & Gaunt, 1985; Gaunt & Gaunt, 1985). This criticism can be answered if the action of the vocal tract filter is responsible for limiting the emitted sound to a single, narrow band of frequencies (Nowicki, 1987).

A second point concerns the coupling between amplitude modulation and frequency modulation discussed earlier. If the acoustic sources alone are responsible for sound generation, then their proposed mode of action must account for this unusual property of birdsong. Indeed, one of the strengths of the Greenewalt model is its ability to explain the coupling based on the behavior of the oscillating MTMs alone (Greenewalt, 1968; Stein, 1968). If, on the other hand, resonance plays a role in production, then an alternative explanation arises. As a frequency modulated signal sweeps from one side of the center frequency of a bandpass filter to the other, the amplitude first waxes as frequency increases, and then wanes, on the other side of the filter center frequency, with further increase in frequency. Thus, one could envision the vocal tract filter actively tracking syringeal output on a syllable-by-syllable basis, but unable to track the most rapid modulations within a single note. This proposal is similar to one raised by Gaunt and Wells (1973) in which mechanical resonances were suggested as the basis of modulation coupling in birdsong.

A final point concerning the nature of the acoustic sources emerges from studies of birds singing in helium (Nowicki, 1987). As with most types of whistle, the generation of a hole-tone whistle requires acoustic feedback to stabilize the rate at which vortices are shed (see Chanaud, 1970, for a general review of the acoustic properties of whistles). In the absence of such stabilization, the acoustic disturbance will be irregular, resulting in a noisy sound instead of a tonal one. In the case of a whistling tea-kettle, for example, the small orifice through which the steam emerges is not a single hole, but in fact two holes on separate plates between which is a small space. The vortices responsible for sound generation are shed within this tiny chamber (i.e., "upstream" from the first hole) and excite the chamber's resonance frequency. The resonance, in turn, affects the initial source of instability, causing the vortex shedding to become regular. Thus, the fundamental frequency of the kettle's whistle is dependent on this resonance.

If birdsong is generated as a hole-tone whistle, or "Rayleigh bird call," then one should observe a significant increase in the frequency of the fundamental in helium because of changes in resonance frequences. The fact that an increase was not observed (Nowicki, 1987) provides a serious difficulty for the "whistle hypothesis" (Nottebohm, 1976; Gaunt et al., 1982). This difficulty may not be insurmountable. Neither of the taxa for which the whistle hypothesis was first proposed (Psittaciformes, or parrots, and Columbiformes, or doves) has been tested in helium. It is also true that the fundamental properties of aerodynamic whistles are still only poorly understood, and it is conceivable that an unusual form of whistle operates in the avian syrinx. In songbirds, however, it seems most likely at present that the acoustic sources are not primarily aerodynamic, but involve the vibration of physical structures such as the medial tympaniform membranes.

VOCAL LEARNING AND VOCAL COORDINATION

The picture emerging from this research is one of a complicated pattern of behavior, involving close coordination between events in the syrinx and configurations of the vocal tract. This coordination is designed to achieve constant readjustment of resonances, with what must often be great speed and precision, to match the changing pattern of syringeal output. Could it be that the ability to operate this complex instrument is instinctive? Current indications are that the skill is more likely to be acquired than innate.

It is now well established that the song patterns of most and perhaps all songbirds are learned, in the sense that songs invariably develop abnormally in birds reared in isolation (Kroodsma, 1982). Local song dialects, a consequence of the dependence of normal vocal development upon auditory experience, are common (Mundinger, 1982; Baker & Cunningham, 1985). At least two types of auditory experiences are relevant. Not only must a young bird be able to hear adult songs in order to learn them, but it is also necessary for a bird to hear its own voice during development. Songbirds typically learn from memory, acquiring songs destined for imitation at one time, and reproducing them at a later time by reference to a stored memory trace or, as it has been termed, an acquired auditory template. Auditory feedback is necessary for this conversion of information in memory storage into an imitated song pattern (Konishi, 1965; Marler & Peters, 1977, 1982). The importance of this role of auditory feedback is revealed by a comparison between the songs of birds reared in isolation with hearing intact and birds which are totally deaf (Figure 16).

Although isolation-reared birds display obvious abnormalities in their song, there also remain many normal features, sufficing to permit a trained observer to distinguish the songs of one species from another (Marler & Sherman, 1985). Basic syntactical features of song often develop within the normal, species-typical range (Figure 16). Such is not the case with songs of deaf birds. Much of the syntactical structure that can be discerned in the song of a hearing bird reared in isolation is now lacking. This contrast led to the postulation of auditory song templates, hypothetical brain mechanisms conceived of as existing in at least two conditions, as innate templates on the one hand, and acquired templates on the other (Konishi, 1965; Marler, 1970; Marler & Sherman, 1985). This line of theorizing has implications for our present theme, throwing some light on the developmental basis of tonal phonology in birdsong.

Deaf birds display a virtually complete loss of note structure. Their songs are often little more than squeaky bursts of noise, so deficient in bird-like qualities that they elicit no response at all when played to wild males (Searcy & Marler, 1987). In many cases songs are so degraded that no species differences remain other than song duration and overall frequency range. Certain species differences in song segmentation can be discerned statistically (Marler & Sherman, 1983), but the variation is great, and many songs are indistinguishable from one species to another. There is no longer any trace of normal song tonality.

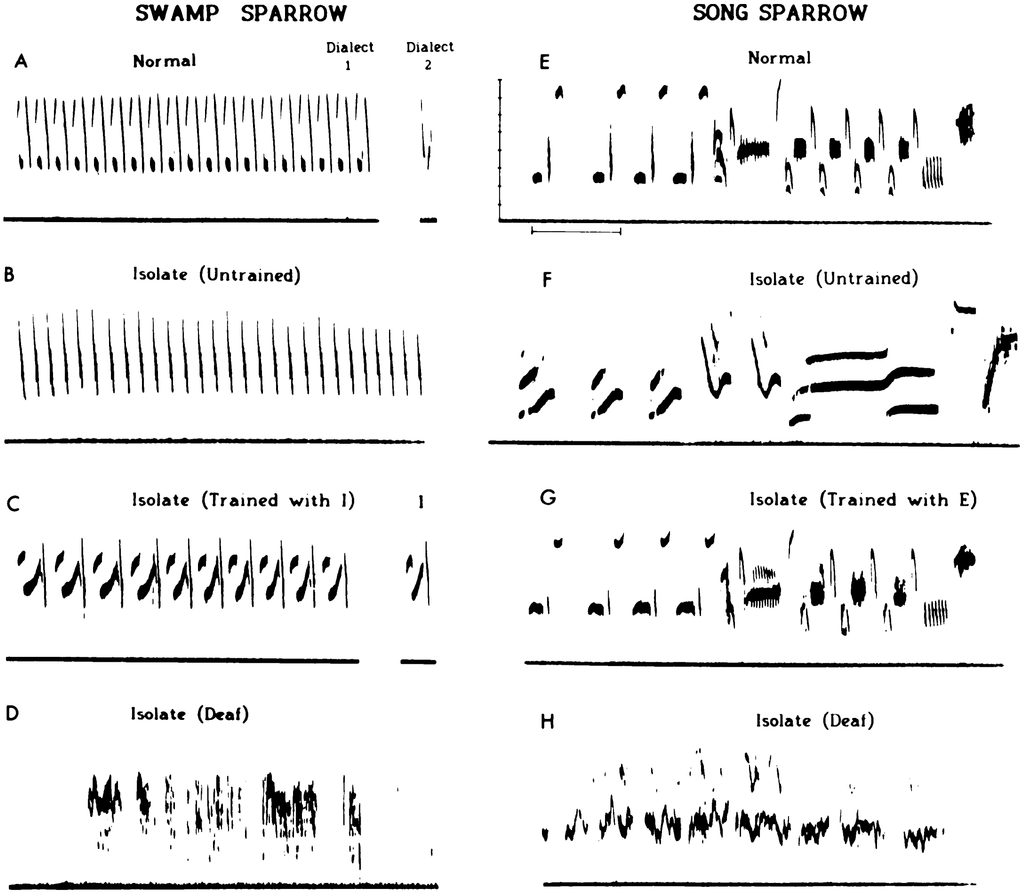


Fig. 16. Sonagrams illustrating typical exemplars of songs produced by song and swamp sparrows under four conditions: (1) normal learning in the wild (A & E, with one syllable of a second dialect also shown in A), (2) acoustic isolation, but with intact hearing (B & F), (3) isolated, but trained with tape recordings of normal adult songs of their species (C & G), and (4) isolated and deafened (D & H). (Kay Digital Sona-Graph Model 7800, 8-kHz analysis range, 300-Hz analysis bandwidth, frequency markers are 1-kHz intervals, time marker is 0.5 sec.)

A comparison between songs of deaf birds and birds deprived of normal adult models, but capable of hearing themselves, reveals some striking differences (Figure 16). In addition to differing species-specific patterns of syntax, hearing birds are also able both to generate tonal sounds and to reiterate song notes in a consistent fashion to produce a trill, something of which deaf birds are incapable. Thus reinstating the auditory feedback loop has radical effects on note structure, restoring many aspects of normal development. Evidently, innate auditory song templates embody some information about how the vocal instrument should be used, made available by their influence on how a song should sound to the singer himself. This information, although not sufficient to completely specify normal adult song, is exploited even if birds are deprived of access to normal models. There are still abnormalities in the tonal quality of songs of isolated birds, however, suggesting that access to external models also influences the development of normal coordination between the two sides of the syrinx, and between syrinx and vocal tract. As can be discerned in Figure 16F, songs of isolates often have a complex spectral structure, with a much stronger emphasis on overtones than is the case in normal song. In fact, there are often striking similarities between the spectral organization of songs of birds reared in isolation and songs of birds singing in a helium atmosphere (see Marler & Sherman, 1985; Nowicki, 1987). This correspondence is so widespread as to suggest that the capacity to develop normal vocal coordination is dependent upon some form of learning, the precise nature of which has yet to be established. It may be a direct consequence of experience of adult songs that are themselves tonal in nature.

Summary and Conclusions

The model we have presented of the ways in which the voice of a songbird operates comes much closer than previous interpretations to the mechanisms that operate in human voice production, perhaps especially as they are thought to be involved in singing. Far from being driven entirely by the vibration of syringeal membranes, independently of resonances of the vocal tract, as previously thought, it now appears that both are involved, functioning in intricate harmony. The very closeness of this coordination between events in the syrinx and the vocal tract, and the rapidity with which adjustments appear to be made may have been one of the obstacles to earlier discovery of this mechanism. This new conception confronts the physiologist with greater challenges than earlier models, requiring as it does cooperation between structures with separate patterns of innervation arising in different parts of the brain.

The "two-voiced" nature of many birdsongs, with melodies that are in some degree separately produced by the two sides of the syrinx (Figure 7), helps to explain the elaborate acoustic structure of birdsong (Greenewalt, 1968), but has also led to misunderstandings, particularly with regard to the supposed independence of the two syringeal sound sources. In some cases there is clear evidence of coupling between them (Nowicki & Capranica, 1986b). The full range of acoustic consequences of this coupling, and the forms it can take, remains to be established. We do not know how widespread this bilateral interaction will prove to be and, indeed, whether there are any cases in which there is totally independent operation of the right and left syringeal sources.

Another issue for the future concerns the way in which the bird brain achieves mastery of this remarkable instrument, producing patterns of sound with qualities that are intriguingly reminiscent in many respects of the sounds of our own music. At least these are tractable problems. Much more difficult to confront, both scientifically and conceptually, are questions about how the birds themselves perceive their songs. Do they have a sense of beauty? When females display a preference for the song of one male over another, are such preferences simply a manifestation of perceptual processes that are automated and reflexive, or is it conceivable that, as Darwin presumed, birds have a primordial aesthetic sense? With the advent of new developments in cognitive science, questions such as these may eventually prove to be accessible to scientific investigation. One hopes that, in the process, scientific convention will retain sufficient flexibility for researchers to admit to the pleasure of listening to birdsongs, even if it should emerge that the birds themselves are unaware of the beauty of the sounds they create.[1]

References

Ames, P. L. The morphology of the syrinx in passerine birds. *Bulletin Peabody Museum of Natural History*, 1971. *37*, 1–194.

Armstrong, E. A. *A study of bird song*. London: Oxford University Press, 1963.

Arnold, A. P., & Bottjer, S. W. Cerebral lateralization in birds. In S. D. Glick (Ed.), *Cerebral lateralization in nonhuman species*. New York: Academic Press, 1985.

Baker, M. C., & Cunningham, M. A. The biology of bird song dialects. *Behavioral and Brain Sciences*, 1985, *8*, 85–133.

Baylis, J. R. Avian vocal mimicry: its function and evolution. In D. E. Kroodsma & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol. 2). New York: Academic Press, 1982.

Beil, R. G. Frequency analysis of vowels produced in a helium-rich atmosphere. *Journal of the Acoustical Society of America*, 1962, *34*, 347–349.

Benade, A. H. *Fundamentals of musical acoustics*. London: Oxford University Press, 1976.

Beranek. L. L. *Acoustic measurements*. New York: Wiley, 1949.

Bergmann, H.-H., & Helb, H.-W. *Stimmen der Vögel Europas*. Munich, Zurich: BLV Verlagsgesellschaft, 1982.

Borror, D. J., & Reese, C. R. The analysis of bird songs by means of a vibralyzer. *Wilson Bulletin*, 1953, *65*, 271–276.

Borror, D. J., & Reese, C. R. Vocal gymnastics in wood thrush songs. *Ohio Journal of Science*, 1956, *56*, 177–182.

[1] Resources for this research were provided by Biomedical Research Support Grant S07 RR07065, USPHS Postdoctoral Fellowship 5 F32 NS07520, and USPHS Grants NS 24651 and MH 14651. Esther Arruza prepared figures and typed the manuscript. We dedicate the paper to the Mary Flagler Cary Trust for their support of research on music, both ornithological and musicological.

Boughey, M. J., & Thompson, N. S. Species specificity and individual variation in the songs of the brown thrasher (*Toxostoma rufum*) and the catbird (*Dumetella carolinensis*). *Behavior*, 1976, *57*, 64–90.

Brackenbury, J. H. Control of sound production in the syrinx of the fowl *Gallus gallus*. *Journal of Experimental Biology*, 1980, *85*, 239–251.

Brackenbury, J. H. The structural basis of voice production and its relationship to sound characteristics. In D. E. Kroodsma & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol. 1). New York: Academic Press, 1982.

Broad, D. T. Phonation. In F. D. Minifie, T. J. Hixson, & F. Williams (Eds.), *Normal aspects of speech, hearing and language*. Englewood Cliffs, NJ: Prentice-Hall, 1973.

Casey, R. M., & Gaunt, A. S. Theoretical models of the avian syrinx. *Journal of Theoretical Biology*, 1985, *116*, 45–64.

Chanaud, R. C. Aerodynamic whistles. *Scientific American*, 1970, *221(1)*, 40–46.

Chanaud, R. C., & Powell, A. Some experiments concerning the hole and ring tone. *Journal of the Acoustical Society of America*, 1965, *37*, 902–911.

Craig, W. The song of the wood peewee *Myiochanes [contopus] virens*. *Bulletin New York State Museum*, 1943, *334*, 1–186.

Cuvier, G. Sur les instruments de la voix des oiseaux. *Journal de Physique, de Chimie, et d'Histoire Naturelle*, 1800, *50*, 426–451.

Cuvier, G. *Leçons d'anatomie comparée*, Vol. 6, 1805.

Darwin, C. *The descent of man and selection in relation to sex*. New York: The Modern Library, 1874.

Dooling, R. J. Auditory perception in birds. In D. E. Kroodsma & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol. 1). New York: Academic Press, 1982.

Dooling, R. J. Perception of complex sounds by birds and humans. In R. J. Dooling & S. Hulse (Eds.), *The comparative psychology of complex acoustic perception*. Hillsdale, NJ: Erlbaum, in press.

Fant, G. *Acoustic theory of speech production*. The Hague: Mouton, 1960.

Gaunt, A. S. Phonation. In T. J. Seller (Ed.), *Bird respiration*. Boca Raton, Fl: CRC Press, 1987.

Gaunt, A. S., & Gaunt, S. L. L. Syringeal structure and avian phonation. In R. F. Johnston (Ed.), *Current ornithology* (Vol. 2). New York: Plenum, 1985.

Gaunt, A. S., Gaunt, S. L. L., & Casey, R. M. Syringeal mechanics reassessed: evidence from *Streptopelia*. *Auk*, 1982, *99*, 474–494.

Gaunt, A. S., & Wells, M. K. Models of syringeal mechanics. *American Zoologist*, 1973, *13*, 1227–1247.

George, J.C., & Berger, A. T. *Avian myology*. New York: Academic Press, 1966.

Greenewalt, C. H. *Bird song: Acoustics and physiology*. Washington: Smithsonian Institution Press, 1968.

Häcker,V. *Der Gesang der Vögel, seine anatomischen und biologischen Grundlagen*. Jena: Gustav Fischer, 1900.

Hall-Craggs, J. The aesthetic content of birdsong. In R. A. Hinde (Ed.), *Bird vocalizations*. Cambridge: Cambridge University Press, 1969.

Hardy, H. C., Telfair, D., & Pielemeier, W. H. The velocity of sound in air. *Journal of the Acoustical Society of America*, 1942, *13*, 226–233.

Hartshorne, C. The relation of bird song to music. *Ibis*, 1958, *100*, 421–445.

Helmholtz, H. von *On the sensations of tone (Translated by Alexander Ellis from 4th German edition, with material added by translator, 1954 reprint edition)*. New York: Dover, 1877.

Hérissant, M. Recherches sur les organes de la voix de quadrupèdes et de celle des oiseaux. *Academie Royal Memoires Scientifiques*, 1753, 279–295.

Holywell, K., & Harvey, G. Helium speech. *Journal of the Acoustical Society of America*, 1964, *36*, 210–211.

Hund, A. *Frequency modulation*. New York: Mc-Graw-Hill, 1942.

Huxley, T. H. *A manual of the anatomy of vertebrated animals*. London, 1871.

King, A. P., & West, M. J. Dissecting cowbird song potency: assessing a song's geographic identity and relative appeal. *Zeitschrift für Tierpsychologie*, 1983, *63*, 37–50.

Klatt, D. H., & Stefanski, R. A. How does a mynah bird imitate speech? *Journal of the Acoustical Society of America*, 1974, *55*, 822–832.

Kloek, G. P., & Casler, C. L. The lung and air sac system of the common grackle. *Auk*, 1972, *89*, 817–825.

Koehler, O. Der Vögelgesang als Vorstufe von Musik und Sprache. *Journal Ornithologie*, 1951, *93*, 3–20.

Konishi, M. The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift für Tierpsychologie*, 1965, *22*, 770–783.

Kroodsma, D. E. Learning and the ontogeny of sound signals in birds. In D. E. Kroodsma & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol. 2). New York: Academic Press, 1982.

Ladefoged, P. *Elements of acoustic phonetics*. Chicago: University of Chicago Press, 1962.

Lemon, R. E. Nervous control of the syrinx in white-throated sparrow (*Zonotrichia albicollis*). *Journal of Zoology*, 1973, *171*, 131–140.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, *74*, 431–461.

Lieberman, P. *Speech physiology and acoustic phonetics*. New York: Macmillan, 1977.

Lieberman, P. *The biology and evolution of language*. Cambridge, Mass.: Harvard University Press, 1984.

Marler, P. Tonal quality of bird sounds. In R. A. Hinde (Ed.), *Bird vocalizations*. Cambridge: Cambridge University Press, 1969.

Marler, P. A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology Monographs*, 1970, *71*, 1–25.

Marler, P. The structure of animal communication sounds. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals*. Berlin: Dahlem Konferenzen, 1977.

Marler, P. Song learning: innate species differences in the learning process. In P. Marler & H. S. Terrace (Eds.), *The biology of learning*. Berlin, Heidelberg, New York, Tokyo: Springer-Verlag, 1984.

Marler, P., & Peters, S. Selective vocal learning in a sparrow. *Science*, 1977, *198*, 519–521.

Marler, P., & Peters, S. Long-term storage of learned birdsongs prior to production. *Animal Behaviour*, 1982, *30*, 479–482.

Marler, P., & Sherman, V. Song structure without auditory feedback: emendations of the auditory template hypothesis. *Journal of Neuroscience*, 1983, *3*, 517–531.

Marler, P., & Sherman, V. Innate differences in singing behaviour of sparrows reared in isolation from adult conspecific songs. *Animal Behaviour*, 1985, *33*, 57–71.

Mathews, F. S. *Field guide to wild birds and their music* (1967 reprint edition). New York: Dover, 1921.

McCasland, J. S. Neuronal control of bird song production. *Journal of Neuroscience*, 1987, *7*, 23–39.

Miskimen, M. Sound production in passerine birds. *Auk*, 1951, *68*, 493–504.

Mundinger, P. C. Microgeographic and macrogeographic variation in acquired vocalizations of birds. In D. E. Kroodsma & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol 2). New York: Academic Press, 1982.

Nottebohm, F. Neural lateralization of vocal control in a passerine bird. I. Song. *Journal of Experimental Zoology*, 1971, *177*, 229–262.

Nottebohm, F. Neural lateralisation of vocal control in a passerine bird. II. Subsong, calls, and a theory of vocal learning. *Journal of Experimental Zoology*, 1972, *179*, 35–49.

Nottebohm, F. Vocal behavior in birds. In D. S. Farner & J. R. King (Eds.), *Avian biology* (Vol. 5). New York: Academic Press, 1975.

Nottebohm, F. Phonation in the orange-winged Amazon parrot, *Amazona amazonica*. *Journal of Comparative Physiology*, 1976, *108A*, 157–170.

Nottebohm, F. Asymmetries in neural control of vocalization in the canary. In S. Harnad, R. W. Doty, L. Godstein, J. Jaynes, & G. Krauthamer (Eds.), *Lateralization of the nervous system*. New York: Academic Press, 1977.

Nottebohm, F. Brain pathways for vocal learning in birds: a review of the first ten years. *Progress in Psychobiology and Physiological Psychology*, 1980, *9*, 85–124.

Nottebohm, F., & Nottebohm, M. E. Left hypoglossal dominance in the control of canary and white-crowned sparrow song. *Journal of Comparative Physiology*, 1976, ***108***, 171–192.

Nowicki, S. Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature*, 1987, ***325***, 53–55.

Nowicki, S., & Capranica, R. R. Bilateral syringeal interaction in the production of an oscine bird sound. *Science*, 1986a, ***231***, 1297–1299.

Nowicki, S., & Capranica, R. R. Bilateral syringeal coupling during phonation of a songbird. *Journal of Neuroscience*, 1986b, ***6***, 3595–3610.

Potter, R. K., Kopp, G. A., & Green, H. C. *Visible speech*. New York: Van Nostrand and Company, 1947.

Rayleigh, J. W. S. *The theory of sound* (1945 reprint of 2nd edition). New York: Dover, 1896.

Réthi, L. Untersuchungen uber die Stimme der Vögel, Abt. III: Anatomie und Physiologie des Menschen und der Tiere sowie Theoretische Medizin. *Sitzungsberichte Acad. Wiss. Wien, Math-naturwiss. Klasse*, 1908, ***117***, 93–109.

Rothenberg, M. Acoustic interaction between the glottal source and the vocal tract. In K. N. Stevens & M. Hirano (Eds.), *Vocal fold physiology*. Tokyo: University of Tokyo Press, 1981.

Rothenberg, M. Cosi Fan Tuti and what it means, or, nonlinear source-tract interaction in the soprano voice and some implications for the definition of vocal efficiency. In T. Baer, C. Sasaki, & K. Harris (Eds.), *Laryngeal function in phonation and respiration*. Boston: College-Hill, 1987a.

Rothenberg, M. The control of air flow during loud soprano singing. *Journal of Voice*, 1987b, ***1***, 262–268.

Rüppell, W. Physiologie und Akustik der Vögelstimme. *Journal Ornithologie*, 1933, ***81***, 433–542.

Saunders, A. A. *Bird song*. New York State Museum Handbook, 1929, ***7***, 1–121.

Saunders, A. A. *Guide to bird songs*. New York: Appleton Century, 1935.

Savart, F. Memoire sur la voix des oiseaux. *Annales de Chimie et de Physique*, 1826, ***32***, 5–24.

Schmidt-Nielsen, K. How birds breathe. *Scientific American*, 1971, ***225***, 72–79.

Searcy, W. A., & Marler, P. Response of sparrows to songs of deaf and isolation-reared males: further evidence for innate auditory templates. *Developmental Psychobiology*, 1987, ***20***, 509–520.

Seashore, E. C. *Psychology of music*. New York: McGraw-Hill, 1938.

Seller, T. J. Unilateral nervous control of the syrinx in Java sparrows (*Padda oryzivora*). *Journal of Comparative Physiology*, 1979, ***129***, 281–288.

Setterwall, C. G. Studier öfver Syrinx hos *Polymyoda passeres*. Dissertation, University of Lund, 1901.

Smith, D. G. The role of the sternotrachealis muscles in birdsong production. *Auk*, 1977, ***94***, 152–155.

Stein, R. D. Modulation in bird sounds. *Auk*, 1968, ***85***, 229–243.

Sundberg, J. The acoustics of the singing voice. *Scientific American*, 1977, ***236*** (3), 82–91.

Thorpe, W. H. Talking birds and the mode of action of the vocal apparatus of birds. *Proceedings Zoological Society, London*, 1959, ***132***, 441–455.

Thorpe, W. H. *Bird song: The biology of vocal communication and expression in birds*. Cambridge: Cambridge University Press, 1961.

Warner, R. W. The anatomy of the syrinx in passerine birds. *Journal of Zoology*, 1972, ***168***, 381–393.

Welty, J. C. *The life of birds* (3rd ed.). Philadelphia:W. B. Saunders, 1982.

Wilson, T. A., Beavers, G. S., DeCoster, M. A. Holger, D. K., & Regenfuss, M. D. Experiments on the fluid mechanics of whistling. *Journal of the Acoustical Society of America*, 1971, ***50***, 366–372.