

Songbirds tune their vocal tract to the fundamental frequency of their song

Tobias Riede^{*†‡}, Roderick A. Suthers^{*}, Neville H. Fletcher[§], and William E. Blevins[¶]

^{*}School of Medicine, Jordan Hall, 1001 East Third Street, Indiana University, Bloomington, IN 47405; [†]Institute for Theoretical Biology, Humboldt-University of Berlin, Invalidenstrasse 43, 10115 Berlin, Germany; [§]Research School of Physical Sciences and Engineering, Australian National University, Canberra 0200, Australia; and [¶]Department of Veterinary Clinical Sciences, Diagnostic Imaging Section, Purdue University, West Lafayette, IN 47907

Communicated by Peter Marler, University of California, Davis, CA, February 14, 2006 (received for review July 28, 2005)

In human speech, the sound generated by the larynx is modified by articulatory movements of the upper vocal tract, which acts as a variable resonant filter concentrating energy near particular frequencies, or formants, essential in speech recognition. Despite its potential importance in vocal communication, little is known about the presence of tunable vocal tract filters in other vertebrates. The tonal quality of much birdsong, in which upper harmonics have relatively little energy, depends on filtering of the vocal source, but the nature of this filter is controversial. Current hypotheses treat the songbird vocal tract as a rigid tube with a resonance that is modulated by the end-correction of a variable beak opening. Through x-ray cinematography of singing birds, we show that birdsong is accompanied by cyclical movements of the hyoid skeleton and changes in the diameter of the cranial end of the esophagus that maintain an inverse relationship between the volume of the oropharyngeal cavity and esophagus and the song's fundamental frequency. A computational acoustic model indicates that this song-related motor pattern tunes the major resonance of the oropharyngeal–esophageal cavity to actively track the song's fundamental frequency.

bioacoustics | hyoid motor pattern | larynx | beak gape | vocal tract filter

Although the pulsating throat of a singing bird is apparent to even the casual observer, its importance in vocalization is unknown. Sound is generated deep in the thorax where the avian vocal organ, the syrinx, is located at the junction between the bronchi and trachea. This syringeal sound is modified by the filter properties of the suprasyringeal vocal tract before it radiates from the beak into the environment as song or call notes, but the anatomical basis of these vocal tract filters and their acoustic properties are not well understood.

The modulation of birdsong by vocal tract resonance was first demonstrated by Nowicki (1), who recorded nine species of songbirds singing in a light gas atmosphere (He–O₂). Because the speed of sound in this gas mixture is nearly twice that in air, resonance frequencies of the vocal tract are shifted upward almost an octave, increasing the level of the second harmonic relative to that of the fundamental frequency (f_0). In addition to providing evidence that the source-filter theory of human speech (2) also applies to birdsong, Nowicki (1) observed that second harmonics, which were apparent in light gas, were suppressed in air regardless of their absolute frequency. This finding indicated that the tuning of the resonance filter is not fixed, but can be adjusted to selectively support different fundamental frequencies.

The search for a tracking resonance filter in the songbird's vocal tract has focused on song-related beak movements. Various investigators (e.g., refs. 3–6) have shown that birdsong is often accompanied by frequency-dependent beak movements in which high fundamental frequencies are associated with a wide beak opening, or gape, compared to that at low fundamentals. The fact that experimentally fixing the gape of a singing bird at a large value increases the relative level of the second harmonic when f_0 is low (refs. 4 and 7, and R.A.S., F. Goller, R. Bermejo,

M. Wild, and P. Zeigler, unpublished data) also implicates the beak as part of the vocal tract filter. It is hypothesized that by varying beak gape a bird alters the end-correction that must be applied to its vocal tract, considered as a rigid stopped tube (the trachea) that is flared at its open end by the beak. A wide beak gape should reduce the effective length of the vocal tract and raise its resonance frequency.

However, some data are inconsistent with the beak gape tracking filter hypothesis. The vocal tract filter in young songbirds learning to sing seems to develop before the appearance of frequency-related beak movements. When juvenile song sparrows (*Melospiza melodia*) begin to sing, their songs initially contain prominent higher harmonics. These harmonics become less prominent and song acquires its adult-like tonality between the mid and late phases of plastic song when the maximum beak gape is still less than ≈ 1 mm and is not coordinated with f_0 . Beak movements do not become coordinated with f_0 until song becomes fixed in its adult form when the song sparrows are ≈ 400 days old, well after tonality appears (8). During late plastic phase of song development, these sparrows must be using something other than beak gape to suppress higher harmonics. In addition, theoretical models (9) suggest the acoustic effect of the maxilla and mandible *per se* is relatively small and limited to high fundamental frequencies.

In other experiments, in which the syrinx of eastern towhees (*Pipilo erythrophthalmus*) was replaced by a calibrated sound source, Nelson *et al.* (10) found that changes in beak gape had no significant effect on the vocal tract resonance for frequencies below ≈ 4 or 5 kHz, but frequencies between ≈ 4 and 7.5 kHz were attenuated by a small gape. These data suggest that beak movements control the attenuation, but not the tuning, of a vocal tract filter that suppresses higher harmonics when gape is small. The data do not support the hypothesis that beak movements cause the vocal tract resonance to track the f_0 . The concept of beak gape as controlling either a band-pass or a low-pass filter that attenuates high frequencies is consistent with Larsen and Dabelsteen's (11) data from the European blackbird (*Turdus merula*) and with data from cardinals indicating that a large gape increases the level of higher harmonics for fundamental frequencies below ≈ 3 or 4 kHz (7).

In the following experiments, we use cineradiography to analyze changes in the configuration of the upper vocal tract between the glottis and the beak in spontaneously singing adult male northern cardinals (*Cardinalis cardinalis*) and a computational model (12) to assess the acoustic significance of these changes. We describe a resonance filter in the avian vocal tract, involving cyclical song-related changes in the volume of the

Conflict of interest statement: No conflicts declared.

Abbreviations: FM, frequency modulation; OEC, oropharyngeal–esophageal cavity; f_0 , fundamental frequency.

[†]To whom correspondence should be sent at the present address: National Center for Voice and Speech, 1101 13th Street, Denver, CO 80204. E-mail: tobiasriede@web.de.

© 2006 by The National Academy of Sciences of the USA

oropharynx and cranial esophagus that track the f_0 and can significantly influence the acoustic properties of vocalizations.

Results

Cardinal song is characterized by frequency modulated (FM) sweeps in which the f_0 may be varied by as much as 2 or 3 octaves between ≈ 1 and 9 kHz. Representative data for three of these syllable types, including one upward sweeping syllable (syllable type 1) and two downward sweeping syllables (types 3 and 4), are shown in Fig. 1.

Song is coordinated with cyclical movements of the hyoid skeleton (Fig. 1 C and D) that produce an air-filled supralaryngeal cavity, the volume of which is inversely related to f_0 (see Movie 1, which is published as supporting information on the PNAS web site). The larynx moves with the hyoid skeleton to which it is attached by extrinsic muscles and ligaments (13–15). The avian hyoid apparatus is not attached to the skull, as it is in mammals, allowing greater freedom of movement and a much larger opening into the oropharyngeal cavity (16). The oropharyngeal cavity is enlarged during low fundamental frequencies by a dorsoventral (Fig. 1 C and D, LV; Syll 1, 7.2 ± 1.3 mm $n = 27$; Syll 3, 6.8 ± 1.9 mm $n = 6$; Syll 4, 6.1 ± 1.9 mm $n = 117$; mean maximum movement \pm SD; $\approx 5\%$ of this movement is due to a backward movement of the vertebral column) and craniocaudal (Fig. 1 C and D, LH; Syll 1, 3.5 ± 1.1 mm $n = 23$; Syll 3, 3.7 ± 1.6 mm $n = 6$; Syll 4, 4.1 ± 2.0 mm $n = 113$; mean maximum movement \pm SD) movement of the basihyoid with a simultaneous lateral movement of the cornua (Fig. 1C Cornua; Syll 1, 6.4 ± 2.1 mm $n = 20$; Syll 3, 6.8 ± 1.8 mm $n = 2$; Syll 4, 8.3 ± 2.6 mm $n = 23$; mean maximum distance between cornuae, each cornua moves laterally half of this distance). When f_0 rises, these movements are reversed, reducing the volume of the oropharyngeal cavity (Fig. 1C). The magnitude of hyoid displacement is inversely correlated with changes in f_0 , regardless of the direction of the FM sweep (Fig. 2). Movements of the beak also tend to covary with f_0 , but are less consistent than those of the hyoid and cornua and seem to be acoustically important over only a limited range of frequencies (9, 10).

The volume of the supralaryngeal cavity is further increased during a low f_0 by an expansion, visible in the x-ray images, of the cranial end of the esophagus, which merges with the oropharyngeal cavity to become a single chamber that is part of the supralaryngeal vocal tract (see Movie 1). The length of the resulting oropharyngeal–esophageal cavity (OEC) can be up to twice that of the oropharyngeal cavity alone. This combined laryngeal–hyoidal and esophageal motor pattern can vary the cavity volume several fold during the course of a syllable.

These cyclical, frequency-dependent changes in the dimensions of the OEC suggest that it has an important acoustic role in song, perhaps acting as a variable filter that tunes the vocal tract resonance to the f_0 of the syrinxal sound source. To gain data for a test of this hypothesis, we constructed a three-dimensional model of the OEC by tracing its outline in successive video frames (Fig. 3 A and B; see Movie 1). These data were used to calculate the volume and geometry of the cavity and beak. The maximum volumes computed from this model agree with those measured from oropharyngeal–esophageal casts that were obtained from four dead birds. Both methods gave a maximum volume of ≈ 2.0 ml during the low frequencies (≈ 2 kHz) at the beginning of syllable 1. The minimum volume near the end of the high frequency upsweep (≈ 5 kHz) that terminates this syllable was ≈ 0.6 ml according to the three-dimensional model.

A computational model based on the dimensions of the cardinal vocal tract (see Supporting Text and Fig. 5, which are published as supporting information on the PNAS web site) was then used to predict the vocal tract resonances (12). Computation shows that the frequency of the oropharyngeal–esophageal resonance can be varied by the bird from below that of the first

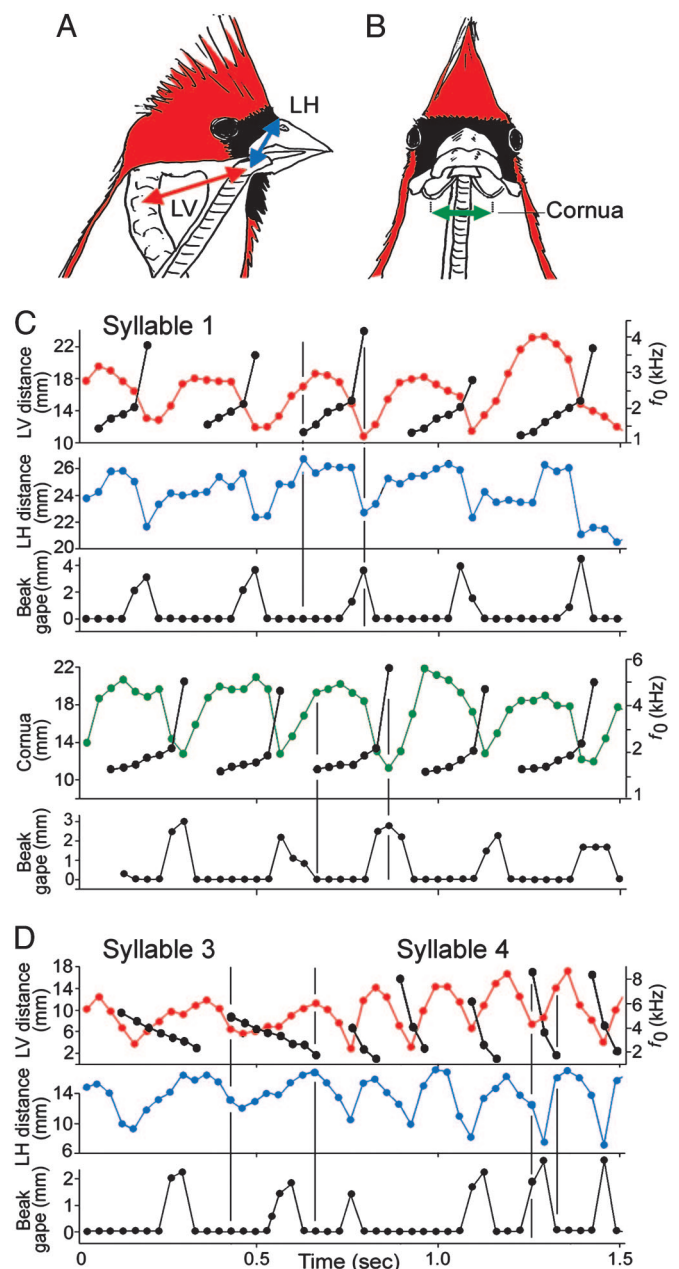


Fig. 1. Each song syllable is accompanied by coordinated movements of the larynx and cornua that maintain an inverse relationship between the size of the OEC and the song's f_0 . (A) Lateral view of cardinal showing the dorsoventral movement (LV) of the larynx from the middle of the second cervical vertebra and its craniocaudal movement (LH) from the dorsal edge of the beak–skull transition. (B) Ventrodorsal view showing distance between lateral cornua of hyoid apparatus (Cornua). (C) Movements of larynx, cornua, and beak during upward sweeping syllable 1. (D) Laryngeal movements during syllables 3 and 4. When f_0 was less than ≈ 2 kHz, beak gape was usually too small to measure on the fluoroscopic images and, although slightly open, was recorded as zero. The peak f_0 of syllables 1 and 4, but not of syllables 2 or 3, was nearly always accompanied by an increased gape. Frequencies of f_0 are superimposed (black) on LV and Cornua distances. LV and LH, syllables 1, 3, and 4 from bird 345; Cornua, syllable 1 from bird 407.

tracheal resonance (≈ 1.6 kHz) to above that of the third tracheal resonance (≈ 8 kHz).

The predicted resonances for syllable type 1 are summarized by the solid curves in Fig. 3C. At the beginning of syllable 1, when the oropharyngeal–esophageal volume is 2.0 ml (Fig. 3A) and

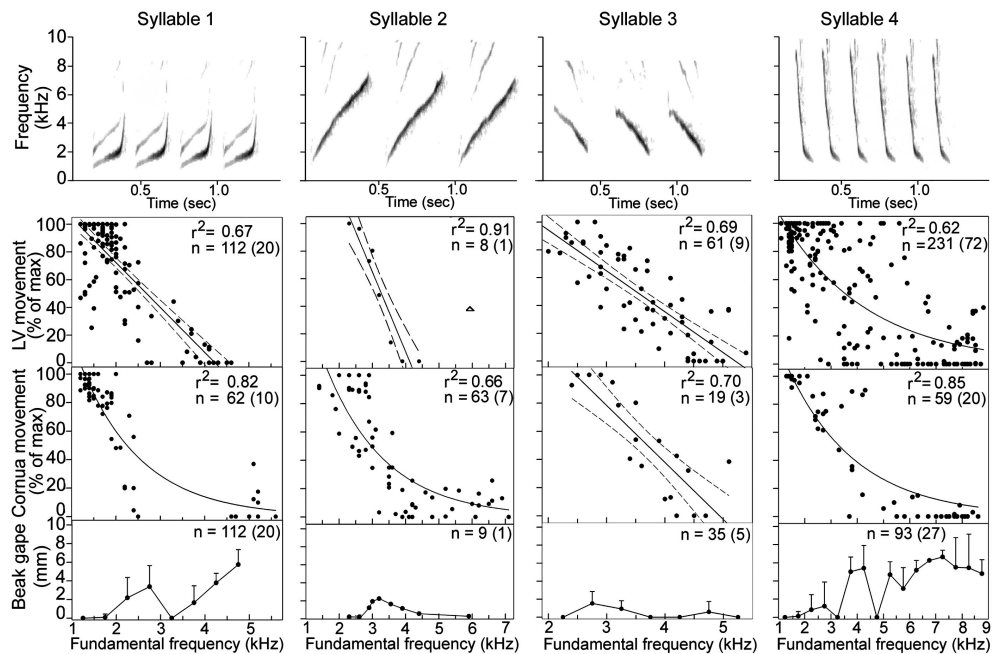


Fig. 2. Relationship between f_0 and dimensions of the supralaryngeal vocal tract for syllables 1–4. Movement of larynx (LV) and cornua (Cornua) is plotted as percent of their maximum movement during the cycle associated with each syllable. Linear regression lines with 95% confidence interval are shown. Data for some relationships are fitted better by exponential regression lines. Beak gape is plotted as mean + 1 SD of values in successive 500-Hz bins. n = number of data points (number of syllables). Each graph is based on a single bird and is consistent with corresponding data from other individuals. The clumping of LV and Cornua data points in the upper left corner of plots for syllable 1 reflects the fact that f_0 was low for most of the syllable's duration so most video frames occurred during this portion of the syllable when hyoid movement was near its maximum. Because of its short duration, syllable 4 was captured by only two or three video frames. Only syllables that contained three video frames are included in the data and this only happened if the first frame coincided with the initial high f_0 and the last frame was near the end of the down-sweep, giving the appearance of a “floor” or “ceiling” effect in some of the plots. The open triangle in LV movement of syllable 2 is not included in the regression line. This data point is from the last video frame of syllable 2, which was followed by a different syllable type starting at 3 kHz. We hypothesize that the bird began to increase the volume of its oropharyngeal cavity before the end of syllable 2, in preparation for the lower initial f_0 of the next syllable.

beak gape is small, the vocal tract has a major resonance peak between ≈ 1.5 and 2.0 kHz (solid purple curve) that coincides with the first tracheal resonance and matches the frequency of f_0 . As f_0 rises, the oropharyngeal–esophageal volume decreases and the major resonance frequency increases, passing through 2.7 kHz (solid dark blue curve), for example, at a volume of ≈ 1.2 ml with a partially open beak. The oropharyngeal volume at the end of syllable 1 is 0.6 ml (Fig. 3B) and beak gape increases to ≈ 6 mm. These changes shift the resonance peak upward, tuning f_0 to ≈ 5 kHz (Fig. 3C, solid light blue curve), which is a little above the second tracheal resonance and in close agreement with the terminal f_0 .

Other syllables are accompanied by similar oropharyngeal and esophageal motor patterns. In some of these f_0 extends as high as 8 or 9 kHz. The bird could tune its vocal tract resonance to these frequencies by reducing the volume of the supralaryngeal vocal tract closer to its collapsed state (Fig. 3C, dashed resonance curves), while maintaining a relatively large beak gape. At high frequencies the cranial end of the esophagus collapses and the volume of the oropharyngeal cavity is reduced as the larynx returns toward its resting position near the base of the beak. An oropharyngeal volume of 0.2 ml, which is still more than twice the volume of the trachea, together with a large beak opening, for example, should shift the major resonance peak to ≈ 9 kHz (Fig. 3C, red curve).

The relationship between the linear dimensions of the oropharynx and f_0 (Fig. 2) and the agreement between f_0 and the predicted resonance of the upper vocal tract at the beginning and end of a FM syllable (Fig. 3) indicate that the first resonance of the oropharyngeal–esophageal filter tracks f_0 , at least during relatively long syllables with gradual FM. A higher sample rate

is needed to determine the accuracy of tracking during the rapidly modulated portion of syllables 1 and 4.

Discussion

The experiments we report here provide several findings that contribute to a better understanding of how songbirds use their vocal tract to modulate their vocal signals. (i) We show that song is accompanied by previously undescribed frequency-dependent changes in the configuration of the upper vocal tract, involving movement of the hyoid skeleton, together with the larynx to which it is attached, and expansion of the cranial end of the esophagus. The larynx of the domestic cock descends during crowing, but its relationship to sound frequency is unknown (17). (ii) This song-related motor pattern maintains an inverse relationship between the volume of the OEC and the song's f_0 . (iii) A computational acoustic model (12) of the upper vocal tract indicates that this motor pattern varies the major resonance of the OEC to track the f_0 , increasing its level relative to that of higher harmonics.

The vocal tract filter of songbirds is fundamentally different from that present in doves, which also expand their upper vocal tract during vocalization. Unlike songbirds, doves vocalize with their mouth and nares closed. They also use expiratory pressure to passively inflate their esophagus, rather than active muscular tension (18). Furthermore, the resonance of the dove's filter is not critically dependent on the volume of the vocal tract, because there is a compensating effect on the physical properties of the esophageal wall (19, 20). However, the most crucial difference is that although sound radiation from songbirds is through the partly open beak, sound radiation in doves occurs by mechanical vibration of the thin walls of the esophagus.

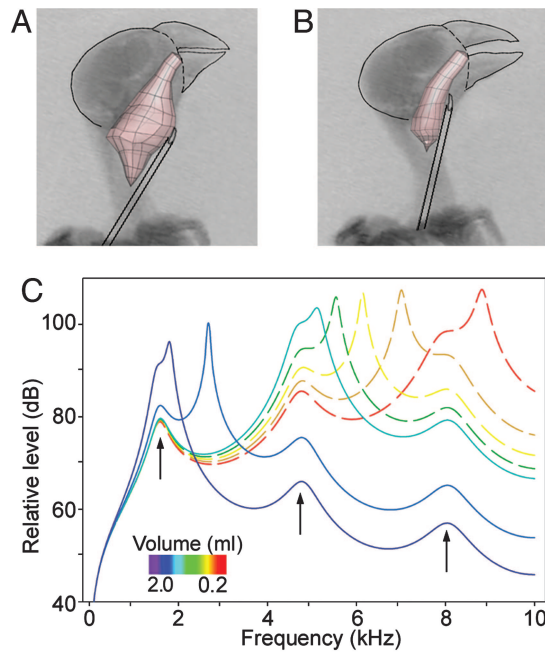


Fig. 3. Three-dimensional reconstructions of the OEC during syllable 1 and predicted vocal tract resonance curves at various vocal tract volumes. (A) At the beginning of syllable 1 (≈ 1.5 kHz), the OEC extends into the cranial end of the esophagus and attains a volume of 2 ml. (B) At the end of syllable 1 (≈ 5 kHz), the esophagus has collapsed and the volume of the cavity is reduced to 0.6 ml. We selected this frequently produced syllable because its long duration provided a more detailed measure, than did shorter syllables, of the relationship between vocal tract shape and f_0 . (C) Predicted resonance curves for the OEC of syllable 1 at volumes of 2 ml and small beak gape (solid purple); 1.2 ml and intermediate beak gape (solid dark blue), and 0.6 ml with wide beak gape (solid light blue). In other syllables, vocal tract resonance (dashed curves) could track f_0 between 5 and 9 kHz by further decreasing cavity volume to 0.5 (green), 0.4, 0.3, or 0.2 (red) ml while holding other parameters, including a wide beak gape, glottal opening, and tracheal length constant. Arrows indicate tracheal resonances.

Combining our findings with those of previous investigators, we conclude that the songbird's vocal tract contains at least three different filter components, the trachea, the OEC, and the beak, that although they interact with each other to some degree, differ in the extent to which they can be tuned and the frequencies over which they are most effective.

The Tracheal Filter. The trachea behaves acoustically like a stopped tube with multiple resonances spaced at intervals across the range of the f_0 (e.g., ref. 21). Tracheal resonances at $\approx 2, 5, 8,$ and 12 kHz, which are close to the frequencies predicted by the computational model, are detectable in FM syllables as a transient increase or decrease in sound level of the odd or even numbered harmonics, respectively, when they cross the predicted tracheal resonances (Fig. 4). The center frequency of the first few of these resonances changes very little during song; this is consistent with measurements indicating that tracheal length changes very little during zebra finch song (22) and with the finding that changes in the beak gape of the eastern towhee have minimal effect on tracheal resonance (10). In theory, the resonances of the songbird tracheal filter could be lowered without changing tracheal length, by constricting the glottis (19). We are unaware of evidence that songbirds normally constrict their glottis during song, but the contribution of the glottis to the vocal tract filter remains to be explored. Tracheal pressure in singing canaries does not increase more than ≈ 1 or 2 cm H_2O above ambient atmospheric pressure (R.A.S., unpublished data), sug-

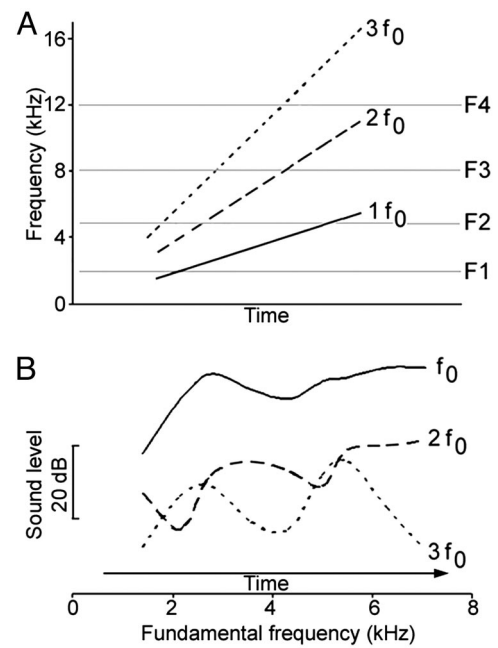


Fig. 4. Sound levels of first, second, and third harmonics in an upward sweeping cardinal syllable. (A) Predicted resonances of a trachea (modeled as a stopped tube 44 mm long) superimposed on a schematic song syllable. (B) Sound level of first, second, and third harmonic, measured every 30 ms during a 310-ms-long FM upsweeping syllable type 2. Assuming the major resonance of the vocal tract tracks the f_0 , these curves plot the peak sound level of a series of resonance curves such as those shown in Fig. 3C. The f_0 is 23 and 33 dB, respectively, above the mean levels of $2f_0$ and $3f_0$. Allowing for assumed differences in source level (see text), this observation suggests the OEC primary resonance may be responsible for ≈ 11 – 14 dB of the sound level difference between the f_0 and its higher harmonics. Tracheal formants are associated with peaks in $3f_0$ near valleys in $2f_0$, as expected for the quarter wavelength resonance of a stopped tube and predicted by the model.

gesting that the glottis is not severely constricted, although small changes in glottal opening might produce acoustic effects before changes in pressure or airflow become apparent.

The Oropharyngeal–Esophageal Filter. The oropharyngeal–esophageal filter is the only one of these three filter components that tracks the fundamental frequency f_0 of the song, perhaps through a reverse mode-locking effect on the syrinxal vibrations, although its accuracy in tracking short steeply FM notes remains to be determined. Its tuning to low frequencies, where it may be particularly important, is made possible by lateral movements of the hyoid cornua, which expand the cranial end of the esophagus, greatly increasing vocal tract volume and allowing the bird to tune its major resonance to frequencies below 2 kHz. At higher f_0 , above ≈ 4 kHz, the esophagus is collapsed, the larynx returns toward its resting position, and beak gape becomes important. It remains to be determined which of ≈ 11 hyoidal and laryngeal muscles are primarily responsible for these movements.

These new song-related peri-oral motor patterns that control the shape of the OEC must be coordinated, by pathways yet to be identified, with the activity of the syrinxal muscles that gate phonation and control the f_0 (23). The computational model we used to calculate the curves in Fig. 3C predicts resonances with quite high Q values (ratio of resonance frequency to bandwidth at -3 dB, ≈ 30) and it could be thought that tuning of the song fundamental to an accurate match might present an insurmountable problem to the bird. There are actually two possible solutions. First, the numerical values used in the model assume

ideally smooth rigid walls, which is clearly an oversimplification. Relaxation of this assumption would add to overall system losses and thus lower the Q value. Second, and perhaps more significantly, the model assumes a simple source-filter system in which the operation of the syringeal valve is not influenced by resonances of the upper vocal tract. An interaction of this sort has not been demonstrated in songbirds, but if present, it could automatically couple the syringeal vibration frequency to the OEC resonance to provide optimal tuning (24).

A direct measure of the contribution of the OEC filter by determining the effect of disabling it is difficult because the muscles involved are also involved in eating and swallowing. An estimate of its importance can be obtained, however, from the song of intact birds. The mean difference between $2f_0$ and f_0 in the radiated spectrum of six cardinal syllables, similar to that in Fig. 4B, was -23 ± 7 dB (M \pm SD). The mean difference between $3f_0$ and f_0 in these same syllables was -34 ± 8 dB. Part of this difference is presumably due to a higher level of f_0 , compared $2f_0$ or $3f_0$, in the signal at the syringeal source. The cardinal's source spectrum is not known, but source spectra of human modal speech decays according to (frequency) $^{-2}$, which equals a relative attenuation of -12 and -19 dB for $2f_0$ and $3f_0$, respectively (25). Estimates of the syringeal source spectrum for the coos of ring doves are in close agreement with those reported for human speech (26). Assuming the source spectrum of a cardinal is subject to a similar decay, the boost in sound level due to the oropharyngeal-esophageal primary resonance may be in the range of ≈ 11 – 15 dB. This is less than the ≈ 20 – 30 -dB difference, predicted by the model, between peak sound levels of the OEC primary resonances and those of tracheal resonances located some distance from that of the OEC (Fig. 3C).

The Effect of Beak Gape. Computational models indicate that the beak has most effect on the filter when it is nearly closed (9). Assessing the specific contribution of beak gape to the vocal tract filter is difficult in living birds because neither the power spectrum of the sound generated by the syrinx nor the extent to which it is modulated by concurrent changes in the configuration of other parts of the suprasyringeal vocal tract are known. Nelson *et al.* (10) attempted to minimize these potential sources of error by measuring the effect of beak gape on a known signal from a miniature speaker inserted into the base of the trachea at the position of the syrinx in dead eastern towhees. In this experiment, opening or closing the beak amplified or attenuated, respectively, frequencies between ≈ 4 and 7.5 kHz by up to 20 or 30 dB, but had little effect on frequencies below ≈ 4 kHz. Thus, the beak seems to exert its major acoustic effect by attenuating frequencies above ≈ 4 kHz, when gape is small. The towhee experiment has the limitation that it eliminates possible feedback or coupling between the oscillating syringeal source and the upper vocal tract. However, it remains to be determined whether source-filter interactions are acoustically important. It is reassuring that data from live cardinals support the conclusion that birds use beak gape to suppress high frequencies. When a singing cardinal is prevented from reducing his beak gape, the attenuation of the second and third harmonics relative to the fundamental is reduced ≈ 20 dB for f_0 below ≈ 3 kHz (R.A.S., F. Goller, R. Bermejo, M. Wild, and P. Zeigler, unpublished data; see ref. 7).

Relationship to Communication. Tonal sounds, i.e., those with relatively little energy in overtones, are common in the songs of many species (27, 28), and there is evidence that birds attend to the tonality of their vocalizations (29–31). The reason for the predominance of tonal sounds in avian communication is not clear. Nelson *et al.* (10) suggest that it may be advantageous to suppress harmonics of fundamental frequencies below 4 kHz that might otherwise degrade the temporal pattern of song elements with an f_0 between 4 and 8 kHz by adding acoustic

“clutter” to this high-frequency portion of the bird's hearing range.

A vocal tract filter that adjusts its major resonance to track f_0 while suppressing higher harmonics may also facilitate acoustic communication by increasing the loudness of the fundamental and uniformity of tone in birdsong. Songbirds thus employ an acoustic strategy similar to that used by human sopranos who adjust the resonance of their vocal tract to match high values of f_0 to be heard over the orchestra (32, 33).

Methods

Cineradiography. X-ray imaging was performed with an x-ray unit having a mobile C-arm (Cardiac Digital Mobile Imaging System; Series 9800). This system provides pulsed digital cine (up to 150 mA) at 30 pulses per second with a 10-ms pulse width and true $1,000 \times 1,000$ imaging resolution, allowing digital recording at 30 frames per s. The digital signal from the fluoroscope was recorded on a S-VHS (NTSC standard) video recorder (JVC Super VHS ET, model HR-S2901U) together with the sound recorded by a directional microphone (Sennheiser ME88) aimed at the bird from a distance of 1 m, as described elsewhere (19). All experiments were in compliance with the guidelines of the National Institutes of Health and were reviewed and approved by the Institutional Animal Care and Use Committee of Indiana University.

Eight different syllable types were recorded from three male northern cardinals as they sang while sitting in the x-ray beam of the C-arm. Data were obtained from all three birds for syllables 1 and 4, and from two birds in the case of syllables 2 and 3. These syllables were selected because they were sung by at least two birds, covered a relatively large frequency range, including both upward and downward sweeps, and provided a reasonable sample size. The data presented here are representative of all birds that sang the syllable. Four syllables were excluded because they had a small sample size from only one bird and/or had a duration < 60 ms and so were represented by only one video frame. Excluded syllables also had a limited bandwidth (< 3 kHz). Based on the available data, hyoid movements in the excluded syllables are consistent with those of the syllables reported here.

Two distances in the lateral (LV and LH) and one in the ventro-dorsal view (Cornua) were measured. LV is the distance between the larynx and the mid-point of the second vertebra. LH is the distance between the larynx and the dorsal edge of the beak-skull transition. Cornua refers to the distance between the most ventral point of the cornua of the hyoid apparatus. Distances were computed (TRACKER, <http://sunflower.bio.indiana.edu/~emartins/tracker.html>) from the coordinates of two points selected manually in each frame. Only cineradiographs with an essentially perfect lateral or ventrodorsal view were used. A reference of known length positioned at the mid-sagittal level of the bird allowed accurate calibration of distance measurements. Ten repeated measures of each of the same three distances in a single frame had a standard deviation of 0.5 mm.

All data are corrected for a delay between the recorded audio and video signals due to the processing time of the x-ray image. The remaining margin of error between the alignment of the vocalization and the x-ray image is estimated not to exceed one-half frame (± 17 ms). Segments of song recorded on the S-VHS video tape were digitized at 30 frames per s (video) and 44.1 kHz (audio) sampling rate (Vegas Video, Sonic Foundry, Madison, WI) and displayed on a computer monitor as individual video frames. Data points were selected with an on-screen cursor. Acoustic measurements were performed by using sound analysis software (PRAAT, version 4.1; www.praat.org).

Computing the Volume of the Supralaryngeal Vocal Tract. The volume of the OEC was estimated by two independent methods. Four male cardinals were killed with an overdose of isoflurane and a cast of the cavity at its maximum volume was obtained by injecting it with impression medium. In two of these birds, threads were attached to the larynx and hyoidal cornua, which were pulled caudo-ventrally or caudo-laterally, respectively, to their positions of maximum displacement during song, before injection of the impression medium. The volume of the cast, or parts of it, was determined by measuring the volume of water it displaced.

The volume of the OEC as it changes during song was also computed from x-ray images of its lateral and ventrodorsal cross-sectional area in each of 10 successive images at 33-ms intervals during the course of a type 1 syllable. Air provides negative contrast in radiological images, making it possible to clearly delineate the borders of the pharynx and the inflated esophagus (see Movie 1). The extents of the oropharyngeal cavity were outlined on both the lateral and ventrodorsal x-ray images using a two-dimensional painting package (PHOTOSHOP, Adobe Systems, San Jose). The outlined lateral and ventrodorsal views were imported into a three-dimensional modeling package

(AC3D, version 4.08; www.ac3d.org), mapped to perpendicular planes, and registered with each other. A three-dimensional polygonal model was then constructed to closely approximate the outlines in both views while also honoring the anatomical constraints of the throat and cavity. Finally, the polygonal model was exported to a three-dimensional geometric analysis tool (RAPIDFORM, 2004 version; www.rapidform.com) to compute the volume of the enclosed space. Maximum OEC volumes computed in this way agreed closely with those obtained from casts. These volumetric data were used in a computational model (12) to predict the resonances of the cardinal's vocal system (see *Supporting Text* and Fig. 5).

We thank Drs. H. Herzel, M. J. Owren, and B. Nelson for their comments on a draft of this article and D. Homberger for discussions on hyoid anatomy. Three-dimensional modeling, volume measurements, and animation were provided by Eric A. Wernert (Advanced Visualization Lab, Indiana University, Bloomington). We are indebted to S. Ronan for technical assistance and to S. A. Zollinger for artwork in Fig. 1. This work was supported by a National Institutes of Health/National Institute of Neurological Disorders and Stroke grant (to R.A.S.) and a fellowship from the Postdoctoral Program of the German Academic Exchange Service (DAAD) (to T.R.).

- Nowicki, S. (1987) *Nature* **325**, 53–55.
- Fant, G. (1970) *Acoustic Theory of Speech Production* (Mouton, The Hague).
- Westneat, M. W., Long, J., John H., Hoese, W. & Nowicki, S. (1993) *J. Exp. Biol.* **182**, 147–171.
- Hoese, W. J., Podos, J., Boetticher, N. C. & Nowicki, S. (2000) *J. Exp. Biol.* **203**, 1845–1855.
- Goller, F., Mallinckrodt, M. J. & Torti, S. D. (2004) *J. Neurobiol.* **59**, 289–303.
- Podos, J., Southall, J. A. & Rossi-Santos, M. R. (2004) *J. Exp. Biol.* **207**, 607–619.
- Suthers, R. A. & Goller, F. (1997) *Curr. Ornithol.* **14**, 235–288.
- Podos, J., Sherer, J. K., Peters, S. & Nowicki, S. (1995) *Anim. Behav.* **50**, 1287–1296.
- Fletcher, N. H. & Tarnopolsky, A. (1999) *J. Acoust. Soc. Am.* **105**, 35–49.
- Nelson, B. S., Beckers, G. J. L. & Suthers, R. A. (2005) *J. Exp. Biol.* **208**, 297–308.
- Larsen, O. L. & Dabelsteen, T. (1990) *Ornis Scand.* **21**, 37–45.
- Fletcher, N. H., Riede, T. & Suthers, R. A. (2006) *J. Acoust. Soc. Am.* **119**, 1005–1011.
- George, J. C. & Berger, A. J. (1966) *Avian Myology* (Academic, New York).
- Homberger, D. G. (1979) *Am. Zool.* **19**, 988.
- Homberger, D. G. (1986) *The Lingual Apparatus of the African Grey Parrot, Psittacus erithacus Linne (Aves: Psittacidae): Description and Theoretical Mechanical Analysis* (American Ornithologist's Union, Washington, DC).
- Homberger, D. G. (1999) in *Proceedings of the 22nd International Ornithological Congress*, eds. Adams, N. J. & Slotow, R. H. (BirdLife South Africa, Durban), pp. 94–113.
- White, S. S. (1968) *J. Anat.* **103**, 390–392.
- Gaunt, A. S., Gaunt, S. L. L. & Casey, R. M. (1982) *Auk* **99**, 474–494.
- Riede, T., Beckers, G. J. L., Blevins, W. & Suthers, R. A. (2004) *J. Exp. Biol.* **207**, 4025–4036.
- Fletcher, N. H., Riede, T., Beckers, G. J. L. & Suthers, R. A. (2004) *J. Acoust. Soc. Am.* **116**, 3750–3756.
- Fitch, W. T. (1999) *J. Zool.* **248**, 31–48.
- Daley, M. & Goller, F. (2004) *J. Neurobiol.* **59**, 319–330.
- Goller, F. & Suthers, R. A. (1996) *J. Neurophysiol.* **76**, 287–300.
- Fletcher, N. H. (1988) *J. Theor. Biol.* **135**, 455–481.
- Flanagan, J. R. (1972) *Speech Analysis Synthesis and Perception* (Springer, New York).
- Beckers, G. J. L., Suthers, R. A. & ten Cate, C. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 7372–7376.
- Greenewalt, C. H. (1968) *Bird Song: Acoustics and Physiology* (Smithsonian Inst. Press, Washington, DC).
- Marler, P. (1969) in *Bird Vocalizations*, ed. Hinde, R. A. (Cambridge Univ. Press, Cambridge, U.K.), pp. 5–27.
- Bremond, J.-C. (1976) *Behaviour* **58**, 99–116.
- Falls, J. B. (1963) in *Proceedings of the XIII International Ornithological Congress*, ed. Sibley, C. G. (American Ornithologist's Union, Ithaca, NY), pp. 259–271.
- Strote, J. & Nowicki, S. (1996) *Behaviour* **133**, 161–172.
- Sundberg, J. (1975) *Acustica* **32**, 89–96.
- Joliveau, E., Smith, J. & Wolfe, J. (2004) *Nature* **427**, 116.