

# Vocal tract motor patterns and resonance during constant frequency song: the white-throated sparrow

Tobias Riede · Roderick A. Suthers

Received: 5 September 2008 / Revised: 23 October 2008 / Accepted: 23 November 2008 / Published online: 10 December 2008  
© Government Employee: National Institutes of Health, USA 2008

**Abstract** Bird song is a complex behavior that requires the coordination of several motor systems. Sound is produced in the syrinx and then modified by the upper vocal tract. Movements of the hyoid skeleton have been shown in the northern cardinal (*Cardinalis cardinalis*) to be extensively involved in forming an oropharyngeal–esophageal cavity (OEC), which contributes a major resonance to the vocal tract transfer function. Here we report that a similar relationship exists between the volume of the OEC and the fundamental frequency in the white-throated sparrow (*Zonotrichia albicollis*) whose song, unlike that of the cardinal, consists of a series of almost constant frequency notes. Cineradiography of singing sparrows shows that the

oropharyngeal cavity and cranial end of the esophagus expand abruptly at the start of each note and maintain a relatively constant volume until the end of the note. Computation of the vocal tract transfer function suggests a major resonance of the OEC follows the fundamental frequency, making sound transmission more efficient. The presence of similar prominent song-related vocal tract motor patterns in two Oscine families suggests that the active control of the vocal tract resonance by varying the volume of the OEC may be widespread in songbirds.

**Keywords** Hyoid skeleton · Vocal tract filter · Songbird · Resonance filter · Gular pump

## Abbreviations

$f_0$	Fundamental frequency
LH	Distance between larynx and beak–skull transition
LV	Distance between the larynx and second vertebra
OEC	Oropharyngeal–esophageal cavity
WTS	White-throated sparrow

**Electronic supplementary material** The online version of this article (doi:10.1007/s00359-008-0397-0) contains supplementary material, which is available to authorized users.

T. Riede  
National Center for Voice and Speech,  
1101 13th Street, Denver, CO 80204, USA

T. Riede  
School of Medicine, Indiana University,  
1001 E. Third Street, Bloomington, IN 47405, USA

R. A. Suthers  
School of Medicine, Department of Biology and Program  
in Neuroscience, Indiana University, 1001 E. Third Street,  
Bloomington, IN 47405, USA

## Present Address:

T. Riede (✉)  
Department of Biology, University of Utah,  
Salt Lake City, UT 84112, USA  
e-mail: t.riede@utah.edu

## Introduction

The acoustic properties of birdsong, like those of human speech, depend not only on the sound produced by the vocal organ, but also on how it is modified by complex song-related motor patterns of the upper vocal tract. By changing the shape of its vocal tract a bird can control the spectral properties of its song in ways that may facilitate vocal communication, but the means by which birds control their vocal tract filter is still a poorly understood aspect of vocal production.

Birdsong is generated in the avian vocal organ, the syrinx, by a pair of labia at the cranial end of each primary bronchus. Each labial pair oscillate, much like the vocal

folds of mammals, when they are adducted by syringeal muscles into the respiratory air stream (Goller and Larsen 1997; Larsen and Goller 1999). Sounds produced in the syrinx are subsequently altered by the suprasyringeal vocal tract filter, which favors the passage of sound near its resonant frequencies (Nowicki 1987; Beckers et al. 2003; Riede et al. 2004; Suthers and Zollinger 2008). Movements of the hyoid skeleton play a major role in tuning of the avian vocal tract filter (Fletcher et al. 2006; Riede et al. 2006).

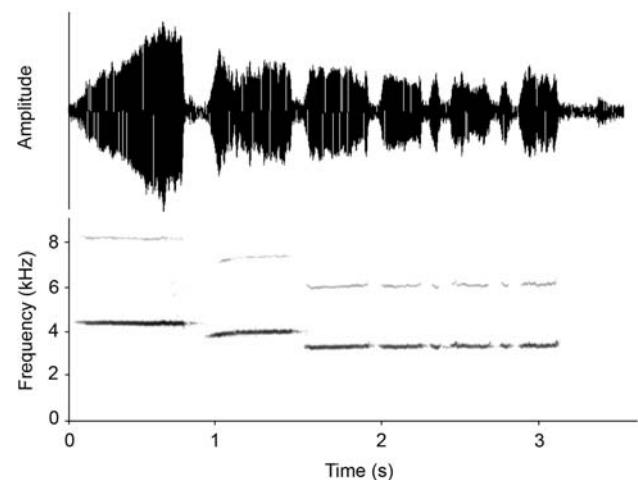
The evolutionary origin of hyoidal control of upper vocal tract movements in birds can be traced back to the respiratory needs of early tetrapods which relied on an oropharyngeal buccal pump to ventilate their lungs. The primitive condition, pure buccal pumping, is found in most air breathing fishes (Brainerd and Owerkowicz 2006), but it also exists as a supplementary breathing pattern during fast locomotion in some reptiles (Owerkowicz et al. 1999). Buccal pumping represents a sophisticated solution for a biomechanical challenge stemming from the fact that muscles only generate force in the direction of shortening. A mechanical construct is necessary that causes an expansion of the oral and pharyngeal cavity in order to generate a negative pressure that sucks in air (or water). The hyoid apparatus provides the lever system that converts muscle shortening into abduction of hyoidal skeletal elements, resulting in an enlargement of the oropharyngeal–esophageal cavity (OEC). The development of thoraco-abdominal aspiration as the primary ventilation system in amniotes removed the respiratory constraints imposed by the buccal pump and allowed the oropharyngeal muscles to assume new tasks. Some of these secondary tasks produce communication signals including visual signals such as the extension of the throat or dewlap in lizards (Font and Rome 1990) and acoustic signals such as the articulation of human speech. The latter involves sophisticated coordination of the hyoid skeleton and hyoidal-pharyngeal muscles in a motor pattern that changes the geometry, and thereby the resonance, of the oral and pharyngeal cavities. Galis (1996) pointed out that structural innovation is essentially associated with four main types of morphological events: acquisition of a new structure; loss of a function by a structure that subsequently becomes available for a new function; duplication of a structure; and functional uncoupling. The use of the hyoid skeleton in acoustic communication in various taxa appears to represent an example where the previous function was lost and a new function evolved.

Cineradiography of singing northern cardinals (*Cardinalis cardinalis*) has shown that they actively control the dimensions of their oropharynx and esophagus so that its primary resonance tracks the fundamental frequency ( $f_0$ ) of the song, even when this frequency is modulated over two or more octaves (Fletcher et al. 2006; Riede et al. 2006). Cardinals increase the volume of the oropharyngeal

cavity and expand the cranial end of the esophagus at low fundamental frequencies, but reduce the volume of these structures when the fundamental frequency of a syllable increases. By maintaining an inverse relationship between the volume of the OEC and the fundamental frequency, the cardinal's OEC helps position and elevate a single vocal tract resonance (Fletcher et al. 2006) tuned to the dominant frequency component generated in the syrinx.

Very little is known about the regulation of the avian hyoid motor pattern during song or its relationship to the song control system (e.g., Wild 1997). The hyoid motor pattern that accompanies song is reminiscent of the complex, learned articulatory maneuvers involved in human speech and thus represents another parallel between the production of birdsong and speech. Articulatory maneuvers in humans are acquired through a learning process (Kuhl et al. 2005; Kuhl and Rivera-Gaxiola 2008) before they are used in speech, posing the possibility that avian hyoidal motor patterns might also involve vocal learning.

Here we report on the hyoid motor pattern and its relationship with fundamental frequency in the white-throated sparrow (WTS; *Zonotrichia albicollis*). We chose the WTS because its song, unlike the strongly frequency modulated songs of the northern cardinal, consists of a series of almost constant frequency whistled notes with relatively little energy in higher harmonics (Fig. 1). The song begins with a few long whistled notes followed by a series of shorter notes. There is a change in the fundamental frequency after the first or second note. In most birds this consists of an upward frequency shift, but in some individuals the shift is downward. The great majority of WTS sing a single song type (Borror and Gunn 1965; Hurly et al. 1990, 1991; Weary and Weisman 1992; Falls and Kopachena 1994).



**Fig. 1** A time-series and spectrogram of a descending song from white-throated sparrow 516. Syllables are nearly constant frequency and contain most of their energy in the fundamental

## Methods

### Subjects

Experiments were performed on four tan morph adult male WTS that were caught as adults. Prior to the experiment, birds which had been kept in an aviary on short days were put on long days (14L:10D) and given subcutaneous implants of testosterone.

### Cineradiography and song recording

X-ray imaging was performed with a Series 9800 Mobile C-Arm and 1 k × 1 k Mobile Neurovascular Work Station (OEC Medical Systems, Inc.). This system provides digital cine at 30 pulses/s with a 10-ms pulse width and true 1,000 × 1,000 imaging resolution, allowing digital recording at 30 frames/s. Songs were recorded from each bird as it sang spontaneously in the X-ray beam while sitting in a cage on a single small perch, about 6 cm long. The C-arm was adjusted so that the bird's head was about 10 cm in front of the image intensifier screen. The probability of obtaining either lateral or frontal views of the bird was increased by orienting the perch so it was either parallel or at a right angle, respectively, to the X-ray beam. The digital signal from the fluoroscope was recorded on a video recorder (Sony GVD-1000 Video Walkman; Mini DV format) together with the sound recorded by a directional microphone (Audio Technica model AT835b) aimed at the bird from a distance of about 0.5 m, as described elsewhere (Riede et al. 2004, 2006).

All data are corrected for a 100-ms 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 ( $\pm 17$  ms).

Segments of song recorded on the Mini DV tape were digitized at 30 frames/s (video) and 44.1 kHz (audio) sampling rate (Vegas Video, Sonic Foundry, Madison, WI, USA) and displayed on a computer monitor as individual X-ray frames with their associated sound track. Data points were selected with an on-screen cursor. Acoustic measurements were performed using sound analysis software (PRAAT, version 4.1; [www.praat.org](http://www.praat.org)). Sound files were high-pass filtered at 500 Hz and low-pass filtered at 20 kHz.

### Vocal tract measurements

The changing shape of the upper vocal tract, due to movement of the hyoid apparatus during song, was analyzed from successive X-ray images taken from lateral and frontal views of the bird. Three songs of the same song type were analyzed from each bird in each of these views, except for WTS 527 for which only two songs were available in the

frontal position. Movement of the hyoid apparatus, and of the larynx to which it is attached, were quantified in each frame of the X-ray movie during song by measuring two distances in the lateral (LV and LH) and one in the frontal view (Cornua) (Fig. 2). LV is the distance between the larynx (basihyoid) 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 (Martins 2001) from the coordinates of two points selected manually in each frame. Only X-ray images with an essentially perfect lateral or frontal view were used. A 1-cm long metal wire 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 0.35 mm. For comparison, the gap between two X-ray opaque metal edges, measured with the same procedure, had a standard deviation of 0.15 mm. The average length of LV, LH and Cornua was calculated for each syllable, using all but the first and the last frame of that syllable. Syllables with a duration <60 ms were excluded from further analysis due to the 30 frames/s sample rate. The Pearson correlation coefficient, *R*, was used to quantify the relationship between anatomical distances (LV, LH and Cornua) and the fundamental frequency.

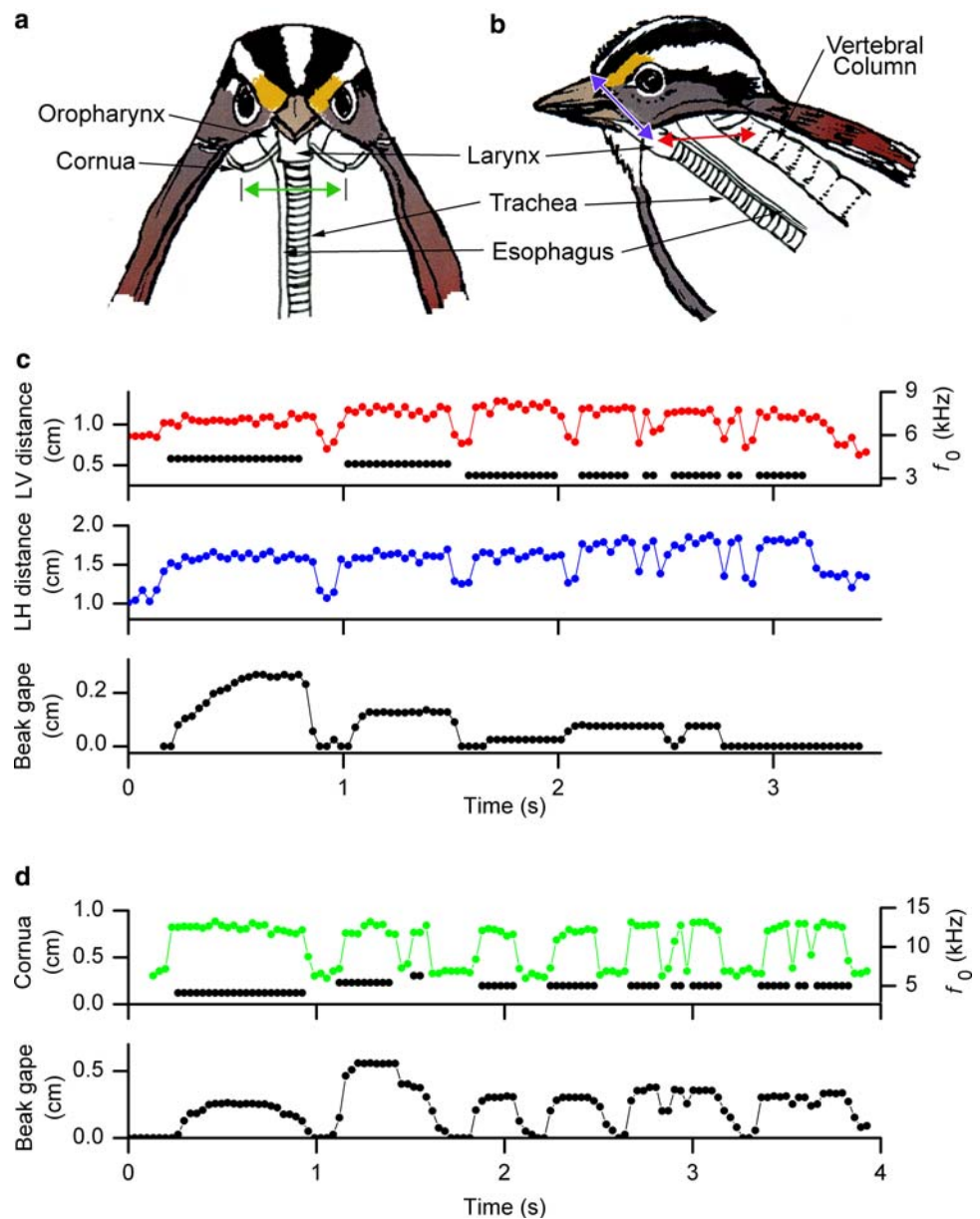
### Beak gape

During most X-ray recording sessions, the bird was also filmed on a digital video camera at 30 fps. Beak gape, the distance between the tip of the maxilla and mandible, was measured from both the X-ray image and the video recording if both were available and of good quality or from only one of these if it was of good quality and the other was not available. Beak gape was excluded from analysis if the X-ray was not clear and the video was not available. Beak gape was measured using the software described above for ‘Video Analysis’.

### Computational model

Fletcher et al. (2006) provided a computational model for calculating the transfer function of a songbird's vocal tract. We measured the dimensions of the trachea in euthanized male WTS. The volume of the OEC was estimated by two independent methods. Three male WTS were euthanized and casts of the suprasyringeal vocal tract, including the OEC cavity, were obtained by injection with dental impression medium. These casts were used to estimate the volume of the maximally expanded OEC by determining the volume of water it displaced in a small graduate cylinder. The volume of the OEC as it changes during song was also

**Fig. 2** Each white-throated sparrow song is accompanied by coordinated movements of the larynx and hyoid cornua that increase the volume of the oropharyngeal–esophageal cavity (OEC) at the beginning of each constant frequency syllable and maintain a relatively constant large OEC volume during the course of each syllable. **a** Schematic frontal view of bird showing the hyoid cornua and the distance measured between them (*green arrow*). **b** Schematic lateral view of upper vocal tract showing dorsoventral movement (LV, *red arrow*) of larynx measured from the middle of the second cervical vertebra and the craniocaudal movement (LH, *blue arrow*) of the larynx relative to the dorsal edge of the beak–skull transition. **c, d** Movements of larynx, cornua and beak during a song are indicated by changes in the distance between the anatomical structures that define the length of each of the vectors measured. **c** LV, LH and beak gape in successive frames of the X-ray movie (30 fps) during a descending song. The timing of each syllable and its fundamental frequency ( $f_0$ ) is plotted in *black* beneath LV. **d** Movement of cornua and beak gape during an ascending song. The fundamental frequency of the song is plotted in *black* beneath Cornua. Artwork in **a** and **b** by S.A. Zollinger



computed from X-ray images. A three-dimensional polygonal model was constructed (modeling package 3D-MAX) to closely approximate the outline of the OEC in frontal and lateral views (see Supplementary Video). The dimensions of these outlines were calibrated with a known reference in the mid-sagittal plane and the volume of the three-dimensional polygon was calculated.

## Results

### Song patterns

Birds 516, 525 and 667 produced descending songs starting with a long syllable of constant  $f_0$  followed by a second

long syllable with a constant but lower  $f_0$  (Fig. 1). Bird 527 sang ascending songs in which  $f_0$  of the second syllable was higher than that of the first (Table 1).

### Movement of the hyoid apparatus and larynx during song

During song, all four WTS exhibited similar characteristic movements of the larynx and the hyoid apparatus to which it is attached. The onset of each syllable was consistently associated with a prominent abrupt increase of all three distances measured (LV, LH, Cornua) (Fig. 2). For the four birds, the larynx moved a mean distance of 4.6 mm ( $n = 117$  syllables) in the ventral direction (LV) and 4.7 mm ( $n = 117$  syllables) in a caudal direction (LH) (Table 2). The lateral movement of the hyoid cornua

**Table 1** Fundamental frequency ( $f_0$ ) and duration of the first three constant frequency syllables in song

Bird ID	First syllable		Second syllable		Third syllable	
	Duration (ms)	$f_0$ (Hz)	Duration (ms)	$f_0$ (Hz)	Duration (ms)	$f_0$ (Hz)
516	613 ± 59	4,383 ± 36	497 ± 38	3,930 ± 21	418 ± 14	3,170 ± 31
525	505 ± 40	5,521 ± 286	430 ± 22	5,465 ± 140	172 ± 17	4,279 ± 88
527	723 ± 52	4,061 ± 44	493 ± 35	5,338 ± 335	216 ± 28	5,043 ± 62
667	866 ± 52	4,275 ± 98	577 ± 48	3,907 ± 61	227 ± 14	3,385 ± 89

Values are mean ± SD.  $n = 10$  Syllables from each bird

**Table 2** Displacement of hyoid measured by three distances during each song syllable

Bird ID	LV (mm)	LH (mm)	Cornua (mm)
516	3.9 ± 1.2; $n = 19$	5.3 ± 0.8; $n = 19$	4.5 ± 0.6; $n = 17$
525	5.1 ± 1.3; $n = 48$	5.8 ± 1.9; $n = 48$	7.3 ± 0.8; $n = 10$
527	3.4 ± 0.6; $n = 7$	2.5 ± 0.4; $n = 7$	4.2 ± 0.6; $n = 19$
667	4.7 ± 1.3; $n = 43$	3.6 ± 1.1; $n = 43$	5.9 ± 1.3; $n = 9$
Average (four birds)	4.6 ± 1.4; $n = 117$	4.7 ± 1.8; $n = 117$	5.1 ± 1.4; $n = 55$

Values are mean ± SD  
 $n$  Number of syllables

increased the mean distance between them by 5.1 mm ( $n = 55$  syllables) (Table 2).

The resulting ventro-caudal displacement of the larynx enlarged the oropharyngeal cavity (birds do not have separate oral and pharyngeal cavities) while the concurrent lateral movement of the cornua expanded the cranial end of the esophagus, resulting in the formation of a single large OEC. This enlarged OEC volume was maintained during the course of the syllable until the hyoid apparatus abruptly returned to its pre-phonatory position at the end of the syllable (see Supplementary Video).

Relationship between movement of hyoid apparatus and fundamental frequency

Quantification of frequency-related changes in the length of the three vectors measured in each X-ray image is difficult for two reasons. First, the frequency difference between notes in the song of individual WTS is small. The  $f_0$  of the lowest and highest notes in the songs of the four WTS we studied differed by only about 0.2–0.3 octave. The difference in the position of the larynx that is needed to adjust the vocal tract resonance to high vs. low frequency notes is therefore also small, especially if the resonance is broadly tuned to the  $f_0$ . Second, given an estimated measurement error of approximately 10% for LH, LV and Cornua due to the small size of the WTS, these small movements may be difficult to quantify.

**Table 3** Correlations between fundamental frequency and distance hyoid apparatus moves

Anatomical distance measured	WTS 516	WTS 525	WTS 527	WTS 667
LV	$R = -0.35$ $P = 0.147$ $n = 19$	<b><math>R = -0.37</math></b> <b><math>P = 0.009</math></b> <b><math>n = 48</math></b>	$R = -0.713$ $P = 0.072$ $n = 7$	<b><math>R = -0.57</math></b> <b><math>P &lt; 0.0001</math></b> <b><math>n = 42</math></b>
LH	<b><math>R = -0.53</math></b> <b><math>P = 0.019</math></b> <b><math>n = 19</math></b>	$R = -0.14$ $P = 0.349$ $n = 48$	$R = -0.006$ $P = 0.099$ $n = 7$	$R = -0.02$ $P = 0.882$ $n = 42$
Cornua	$R = -0.46$ $P = 0.063$ $n = 17$	$R = 0.36$ $P = 0.504$ $n = 10$	$R = -0.09$ $P = 0.719$ $n = 19$	$R = -0.52$ $P = 0.228$ $n = 7$

Average distance change per syllable, standardized to the whole song in which the syllable occurred. Significance level  $P \leq 0.05$

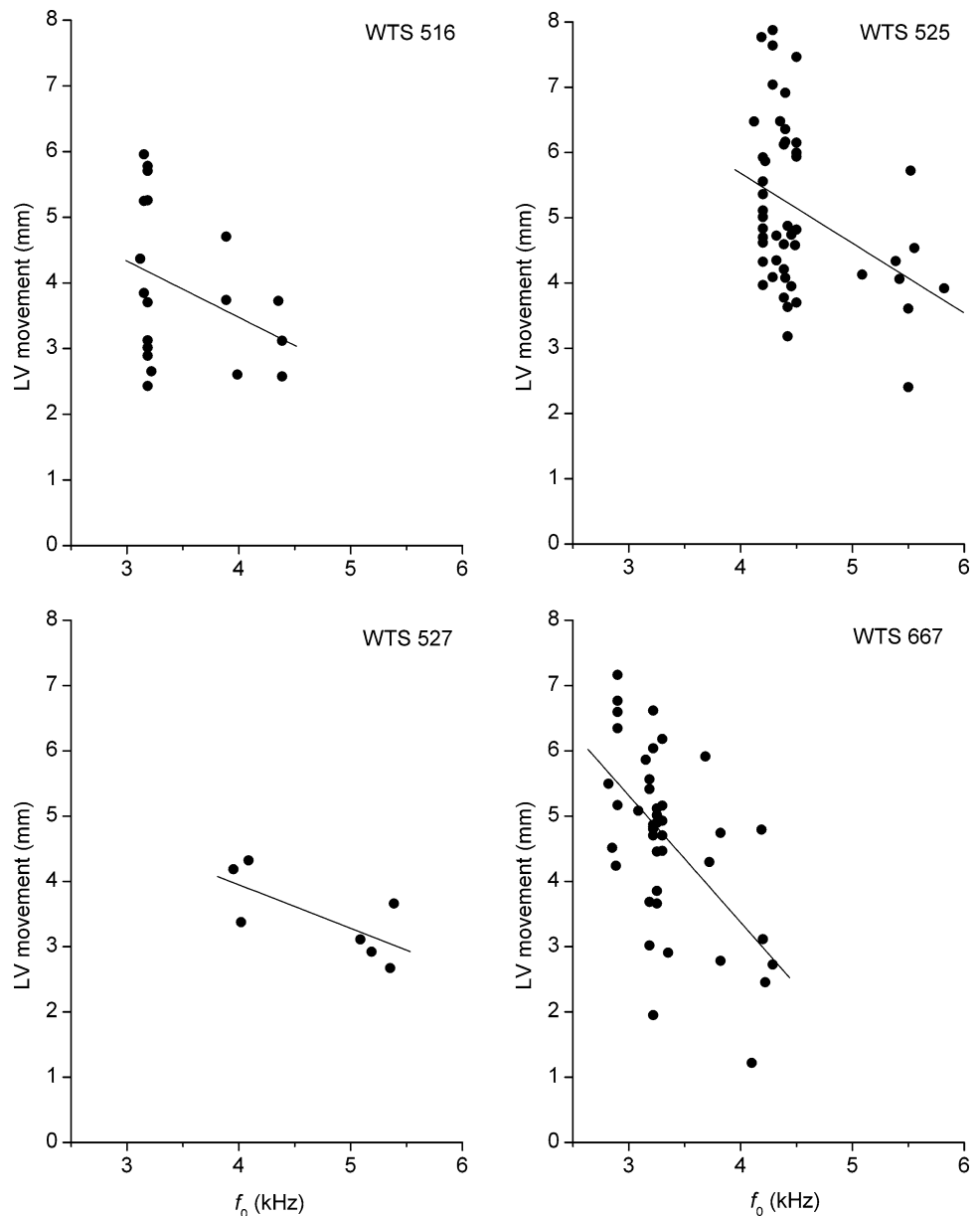
$n$  Number of syllables

The Pearson correlation coefficients between fundamental frequency and each of the three distances measured in X-ray images are summarized in Table 3 and Fig. 3. The linear regression between LV or LH and  $f_0$  is negative in all four birds, as expected assuming the bird maintains an inverse relationship between the volume of the OEC and  $f_0$ . The regression of LV on  $f_0$  is significant in two birds and close to significance in a third (WTS 527). The regression of LH on  $f_0$  is significant in only one bird. The regression for lateral movement of the cornua on  $f_0$  is not significant in any of the birds.

Beak movements

We measured the beak gape and  $f_0$  in each video frame of all syllables in six songs from each bird. Prominent beak movements accompanied most of these syllables, except in the case of WTS 516 which did not open his beak more than a few tenths of a millimeter during most of his syllables. Of the four birds, only WTS 525 had a significant correlation between the mean beak gape and the fundamental frequency of the syllable. In all four birds, the absolute magnitude of the mean gape varied greatly between different syllables having the same constant  $f_0$  (Fig. 4).

**Fig. 3** The movement of LV during song syllables tends to be inversely correlated with the syllable's fundamental frequency. Increasing the distance the larynx moves from the vertebral column during song increases the volume of the oropharyngeal cavity. Each data point represents the mean length of LV during one syllable minus its mean length during all silent intervals between syllables in the same song. Linear regression for WTS 516:  $R = -0.35 \pm 0.12$ ,  $P = 0.147$ ,  $n = 19$ ; WTS 525:  $R = -0.37 \pm 0.12$ ,  $n = 48$ ,  $P = 0.009$ ; WTS 527:  $R = -0.71 \pm 0.05$ ,  $P = 0.07$ ,  $n = 7$ ; WTS 667:  $R = -0.57 \pm 0.11$ ,  $P < 0.0001$ ,  $n = 42$



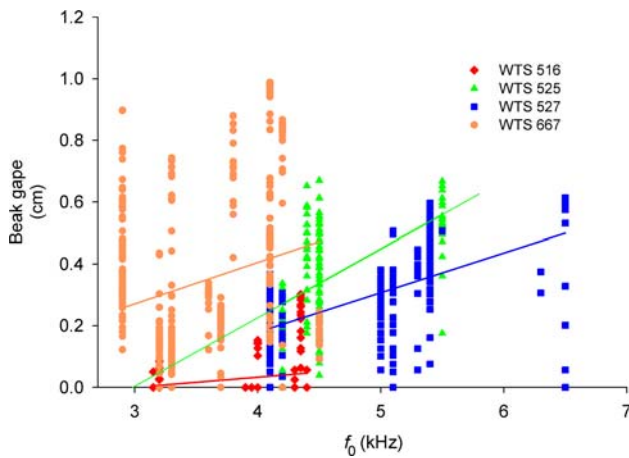
#### Estimation of suprasyringeal vocal tract resonance

Measurements from casts of the vocal tract gave an average maximum OEC volume for four males of  $0.5 \pm 0.1$  ml. The volume of the OEC during song was also estimated by reconstructing its shape from its outline in each successive frame of the X-ray movie during a syllable. The mean estimated OEC volume for eight syllables, including a high and a low frequency syllable in a song from each of four male WTS, was  $0.65 \pm 0.39$  ml (Figs. 5, 6). The regression line for these data indicates that the volume of the OEC has a significant inverse correlation ( $R^2 = 0.53$ ;  $P < 0.05$ ) with the  $f_0$  of the constant frequency syllable (Fig. 6). The OEC tends to be larger

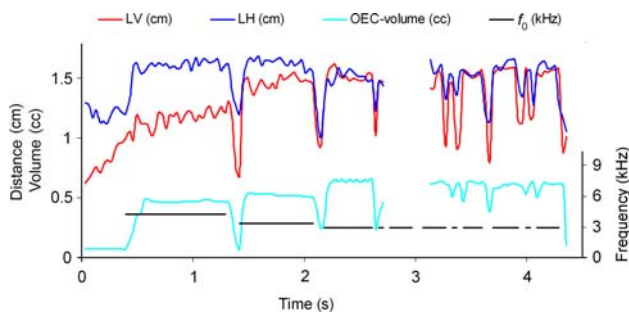
during syllables with a low  $f_0$  than it is during syllables with a high  $f_0$ .

We used a computational acoustic model of the songbird vocal tract (Fletcher et al. 2006) to estimate the resonance curves of the suprasyringeal vocal tract for OEC volumes between 0.2 and 1.4 ml, corresponding to the range of OEC volumes calculated for the eight syllables in Fig. 7. The model predicts a suprasyringeal vocal tract transfer function that is dominated by a major resonance peak that can be tuned to any frequency between about 3 and 8 kHz by adjusting the volume of the OEC to an appropriate value within its experimentally observed range from about 1.4 to 0.2 ml, respectively (Fig. 7).

The first tracheal resonance at about 2.2 kHz was about 20 dB below the OEC resonance. A second tracheal

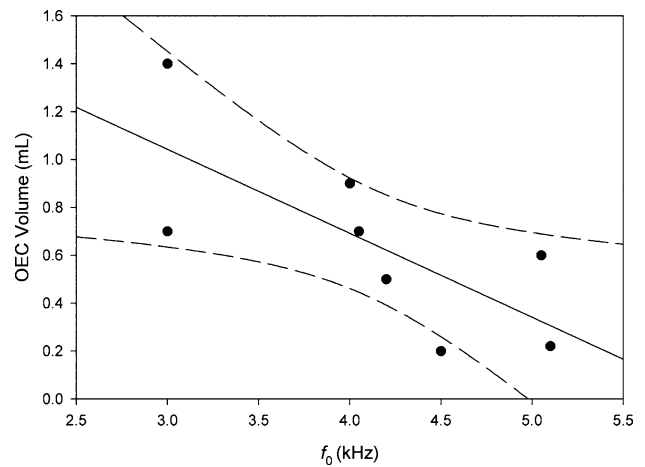


**Fig. 4** Relationship between the mean beak gape and fundamental frequency of constant frequency WTS song syllables. In many cases mean beak gape varied over a wide range for the same  $f_0$ . Beak gape and  $f_0$  were measured in each video frame (30 fps) and the mean gape and  $f_0$  were calculated for each syllable in six songs from each bird. WTS 516:  $R = 0.277$ ,  $P > 0.05$ ,  $n = 47$  syllables; WTS 525:  $R = 0.501$ ,  $P < 0.01$ ,  $n = 26$  syllables; WTS 527:  $R = 0.303$ ,  $P > 0.05$ ,  $n = 32$  syllables; WTS 667:  $R = 0.011$ ,  $P > 0.05$ ,  $n = 69$  syllables. Frequency bins are 50 Hz wide. Beak gapes less than about 0.2 mm could not be accurately measured and were assigned a value of 0. This was the case for the majority of WTS 516 syllables, so that data points at zero gape include multiple superimposed syllables

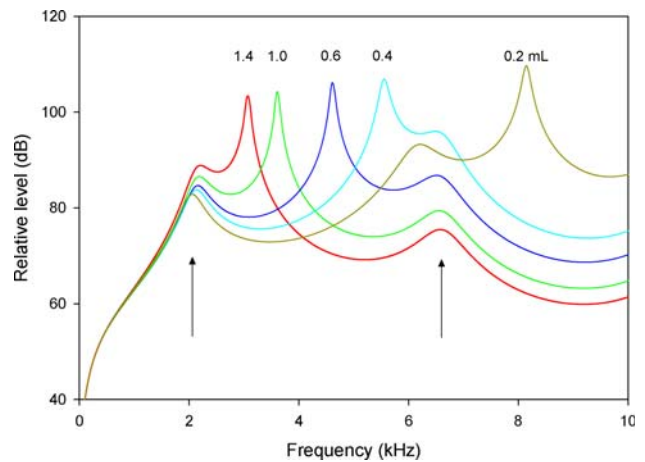


**Fig. 5** The white-throated sparrow typically maintains a nearly constant OEC volume during each constant frequency syllable. In this song from WTS 667, the OEC volume is inversely related to the syllable's  $f_0$ . The increased volume is due in part to a ventro-caudal movement of the larynx, indicated by the increase in distances LV and LH, that enlarges the oropharynx and by expansion of the cranial end of the esophagus (not shown). The volume of the OEC was calculated from its outline in X-ray images, not from LV or LH. The absence of morphological data at about 3 s is due to the bird briefly turning its head

resonance at about 6.6 kHz was also present. The frequency of tracheal resonances was not significantly affected by changes in the volume of the OEC. The length of the trachea in three freshly dissected, male WTS was 3.4 cm. In these same males, the cast of the trachea, made by injecting dental impression medium into the vocal tract through the mouth, was 3.8 cm long. The 4 mm increase in the length of the cast compared to the trachea before the injection of the impression medium, suggests a longitudinal elasticity in



**Fig. 6** Inverse relationship between the volume of the oropharyngeal–esophageal cavity and the fundamental frequency. Data points include the highest and lowest constant frequency syllables from each of four white-throated sparrows.  $R^2 = 0.53$ ;  $P = 0.040$ ; dashed lines indicate 95% confidence interval



**Fig. 7** Predicted resonance curves of the entire suprasyringeal vocal tract for five different volumes of the oropharyngeal–esophageal cavity. By varying only the volume of the oropharyngeal–esophageal cavity between 0.2 and 1.4 ml a white-throated sparrow could adjust the principal resonance peak of its vocal tract to match the fundamental frequency of its constant frequency syllables, thus significantly increasing the amplitude of the song. Arrows indicate predicted resonances of a 36-mm long trachea closed at the syrinx and open at the glottis. The length of the WTS beak is about 11 mm, however the effective length of the beak decreases with increasing gape. The resonance curves in this figure assume an effective beak length of 2.3 mm, which corresponds to a relatively large beak gape. Based on the Fletcher et al. (2006) computational model of the avian vocal tract

the trachea that may allow WTS to vary the length of their trachea by about 11%. For all resonance curves we assumed a tracheal length of 36 mm and an effective acoustic length for the beak of 2.3 mm, which is equal to about 20% of its geometrical length and corresponds to a relatively wide beak gape (Fletcher et al. 2006).

## Discussion

### The oropharyngeal–esophageal motor pattern

In previous experiments we (Riede et al. 2006) have shown that the northern cardinal, whose song contains prominent frequency modulated sweeps, maintains an inverse relationship between the volume of its OEC and the changing fundamental frequency of its song. A computational model of the cardinal vocal tract (Fletcher et al. 2006) indicates that singing cardinals actively adjust the volume of their OEC so that its primary resonance tracks the  $f_0$  of the song.

In this paper we document similar song-related changes in the dimensions of the oropharynx and esophagus of the WTS, a species that is not closely related to the cardinal and whose song consists of constant frequency syllables with almost no frequency modulation. We show that the onset of each of these WTS syllables is accompanied by a rapid ventro-caudal movement of the hyoid apparatus and larynx together with a lateral movement of the hyoid cornua. This movement enlarges the oropharyngeal cavity and cervical esophagus which are maintained in their enlarged state with little variation until the end of the syllable when the larynx and hyoid apparatus return to their pre-phonatory position and the OEC collapses. Most syllables are also accompanied by opening the beak.

### Vocal tract acoustics

#### *The source-filter theory*

According to the linear source-filter theory of human vowel production (Fant 1960), the time-varying glottal airflow at the sound source is not affected by the resonance filter of the upper vocal tract. The independence between the dynamics of the syringeal source and vocal tract resonance filter is an unresolved issue (Nowicki 1987; Beckers et al. 2003; Elemans et al. 2007). However, the formant pattern of vocalizations is determined by the vocal tract as a whole, including the trachea, glottis, OEC and beak, acting as a resonant system.

#### *The trachea and glottis*

The first and second tracheal resonances are close to those predicted for a simple tube closed at the syrinx and open at the larynx. Modeling the trachea as a stopped tube, i.e., as a quarter-wave resonator, predicts a first resonance (in moist air at 41°C) at about 2.6 and 2.3 kHz for lengths of 34 and 38 mm, respectively. The second tracheal resonance is predicted to be about 7.7 or 6.9 kHz, respectively, for these two tracheal lengths. Since the  $f_0$  of syllables sung by our four subjects ranged from about 3.1 to 5.5 kHz (Table 2),

our calculations suggest that the  $f_0$  of most WTS song lies between the first two tracheal formants, where it is not supported by the trachea's resonances.

Little is known regarding the ability of songbirds to vary the length of their trachea. Daley and Goller (2004) determined that zebra finches can only change the length of their trachea a maximum of 3%, equivalent to 1 mm or about a 78-Hz change in the predicted resonance of an equivalent stopped tube. In WTS, the difference between the length of the trachea before and after filling it with impression medium, suggests it may be more elastic than that of the zebra finch.

There are various ways in which a WTS might, in theory, adjust its tracheal resonance to match its  $f_0$ , but they have not been investigated in WTS. Tracheal length might be altered by contraction of the tracheolateralis or sternotrachealis muscles; by activity of hyoidal muscles that move the larynx; or by postural changes during song, such as tilting the head back. Nevertheless, large changes in tracheal length seem anatomically improbable.

Tracheal resonances might also be varied by controlling the glottal aperture. Constricting the glottis should lower the tracheal resonance. In the preceding calculations we assume a glottal opening 1.5 mm in diameter (tracheal diameter averages 1.6 mm). Our model predicts that reducing the glottal diameter to 0.5 mm will lower the second tracheal formant from about 6.9 to 5.0 kHz, which could coincide with the  $f_0$  of WTS syllables in the high frequency portion of the species' vocal range. Glottal aperture during phonation has not been measured in songbirds.

#### *The oropharyngeal–esophageal cavity*

Several kinds of evidence support the hypothesis that the prominent hyoidal motor pattern we observed in WTS tunes the vocal tract's primary resonant frequency to match the  $f_0$ , which is the dominant frequency component generated in the syrinx.

Our cineradiography data show that each syllable is accompanied by an enlarged OEC which is maintained at an essentially constant volume for the duration of each constant frequency syllable, as expected if the sparrow is attempting to make the resonance of its vocal tract match the constant frequency of its song.

If the WTS adjusts the volume of its OEC to resonate at the  $f_0$ , then one would expect an inverse relationship between the ventro-caudal displacement of the larynx (i.e., LV and LH) and the  $f_0$  of each syllable. Since the hyoid cornua appear to control the enlargement of the esophagus, the amplitude of their movement (Cornua) should also be inversely correlated with  $f_0$ . Although in all birds, the onset of each syllable was accompanied by changes in LV and LH in directions that enlarged the OEC, the amount of



movement showed a significant inverse correlation with  $f_0$  for LV in only two birds and for LH in only one bird.

The cornua of the hyoid apparatus always moved laterally at the onset of phonation, but the amplitude of this lateral movement was not significantly correlated with  $f_0$  in any of the birds. It may be that modulation of OEC resonance for the high and low frequency syllables within the song depends on laryngeal movements that do not include the cervical esophagus. If so, it is in contrast to northern cardinals in which lateral movement of the hyoid cornua is related to the expansion of the cranial end of the esophagus (Riede et al. 2006).

Differences in the mean  $f_0$  between high and low frequency syllables in individual WTS songs are relatively small, however, and require correspondingly small amplitude adjustments in the hyoidal motor program. The absence of a consistently significant inverse relationship between each of these hyoid movements and  $f_0$  may be due to limitations in the accuracy with which we can measure these distances in X-ray images of this small bird. If the vocal tract resonance of WTS is broadly tuned, i.e., has a low  $Q$ -value, there may be little advantage in adjusting OEC dimensions for small (<0.5 octave) differences between the  $f_0$  of different syllables within a song. This interpretation is supported by the fact that when the range of  $f_0$  and OEC volumes is increased, by including a song from each bird, there is a significant negative relationship between these variables.

Finally, the resonance curves based on our computational model of the avian vocal tract (Fletcher et al. 2006) (Fig. 7), predict that the range of OEC volumes, which we measured in our four subjects (0.2–1.4 ml; Fig. 6), should have their primary resonance peaks close to the bird's  $f_0$  (5.5–3.1 kHz).

### The beak

Sound in the OEC must pass through the beak before it is broadcast into the external environment. Westneat et al. (1993) showed that the fundamental frequency of the first three notes in WTS song is positively correlated with beak gape. We observed a similar correlation between beak gape and  $f_0$  in many of the syllables we recorded although during some songs (including a majority of those sung by WTS 516) the beak gape was <1 mm. Particularly striking, is the lack of a correlation between absolute gape and  $f_0$  across individual WTS. For example, the beak gape of WTS 667 for  $f_0$  between about 2.9 and 4.2 kHz often exceeded the maximum mean gape of the other three birds, including that of WTS 527 at a  $f_0$  as high as 6.5 kHz. It is interesting that although the regression of beak gape on  $f_0$  suggests a positive relationship, the absolute values of gape as a function of frequency differ for each of the four birds. A 4-mm gape on

the least squares regression line occurs at about 4.0 kHz for WTS 667, 4.8 kHz for WTS 525 and 5.8 kHz for WTS 527.

The acoustic contribution of beak gape to vocal tract resonance is not well understood. On the one hand, several studies of various species have shown an inverse relationship between beak gape and  $f_0$  over at least part of the bird's frequency range (Westneat et al. 1993; Moriyama and Okanoya 1996; Suthers and Goller 1997; Hoese et al. 2000; Goller et al. 2004; Podos et al. 2004). However, physical and computational models of the avian vocal tract (Fletcher and Tarnopolsky 1999; Fletcher et al. 2006) predict that the beak's effective acoustic length rapidly decreases as gape increases from 0 to about 1 or 2 mm, but fail to explain why much larger gapes are typical of most species studied. Some experimental data (Suthers and Goller 1997; Goller et al. 2004; Nelson et al. 2005; Riede et al. 2006) suggest the filter effect of the beak per se is relatively small and limited to high fundamental frequencies. Further research is needed to resolve this dichotomy.

Taken together, our data indicate that WTS, like northern cardinals, adjust the dimensions of their oropharynx and esophagus during song in order to maintain a vocal tract resonance close to the  $f_0$  they are generating in their syrinx. In doing this they, like cardinals, increase efficiency with which the dominant frequency component produced in the syrinx is transmitted through the suprasyringeal vocal tract. This in turn increases its level in the song. Higher harmonics of  $f_0$  that are not close to a resonance peak will undergo relatively more attenuation in the vocal tract, increasing the tonal purity of the song.

**Acknowledgments** We thank Dr. Donna Maney for providing the birds used in this experiment and Amy Coy for assistance in preparation of the figures. Supported by NIH-NINDS R01 NS029467 to R.A. Suthers and by a Postdoctoral fellowship from the 'Deutsche Akademie der Naturforscher Leopoldina' (BMBF-LPD 9901/8-127) to T. Riede. We thank Chauncey Frend and Jeff Rogers (Advanced Visualization Lab, Indiana University, Bloomington) for three-dimensional modeling and the video animation. The experiments reported in this paper were reviewed and approved by the Institutional Animal Care and Use Committee and the Radiation Safety Office of Indiana University, and comply with the 'Principles of animal care', publication no. 86-23, revised 1985 of the National Institute of Health.

### References

- Beckers GJL, Suthers RA, ten Cate C (2003) Pure-tone birdsong by resonance filtering of harmonic overtones. *Proc Natl Acad Sci USA* 100:7372–7376
- Borror DJ, Gunn WWH (1965) Variation in white-throated sparrow songs. *Auk* 82:26–47
- Brainerd EL, Owerkowicz T (2006) Functional morphology and evolution of aspiration breathing in tetrapods. *Respir Physiol Neurobiol* 154:73–88
- Daley M, Goller F (2004) Tracheal length changes during zebra finch song and their possible role in upper vocal tract filtering. *J Neurobiol* 59:319–330

- Elemans CPH, Laje R, Mindlin GB, Goller F (2007) Coupling and source-filter interaction in birdsong. In: Eighth international congress of neuroethology. University of British Columbia, Vancouver, p 97
- Falls JB, Kopachena JG (1994) White-throated sparrow (*Zonotrichia albicollis*). In: Poole A (ed) Birds of North America online. Cornell Laboratory of Ornithology, Ithaca. Retrieved from the Birds of North America online: <http://bna.birds.cornell.edu/bna/species/128>
- Fant G (1960) Acoustic theory of speech production. Mouton, The Hague
- Fletcher NH, Tarnopolsky A (1999) Acoustics of the avian vocal tract. *J Acoust Soc Am* 105:35–49
- Fletcher NH, Riede T, Suthers RA (2006) Model for vocalization by a bird with distensible vocal cavity and open beak. *J Acoust Soc Am* 119:1005–1011
- Font E, Rome LC (1990) Functional morphology of dewlap extension in the lizard (*Anolis equestris*; Iguanidae). *J Morphol* 206:245–258
- Galis F (1996) The application of functional morphology to evolutionary studies. *Trends Ecol Evol* 11:124–129
- Goller F, Larsen ON (1997) A new mechanism of sound generation in songbirds. *Proc Natl Acad Sci USA* 94:14787–14791
- Goller F, Mallinckrodt MJ, Torti SD (2004) Beak gape dynamics during song in the zebra finch. *J Neurobiol* 59:289–303
- Hoese WJ, Podos J, Boetticher NC, Nowicki S (2000) Vocal tract function in birdsong production: experimental manipulation of beak movements. *J Exp Biol* 203:1845–1855
- Hurly TA, Ratcliffe L, Weisman R (1990) Relative pitch recognition in white-throated sparrows, *Zonotrichia albicollis*. *Anim Behav* 40:176–181
- Hurly TA, Weisman RG, Ratcliffe L, Johnsrude IS (1991) Absolute and relative pitch production in the song of the white-throated sparrow *Zonotrichia albicollis*. *Bioacoustics* 3:81–92
- Kuhl PK, Rivera-Gaxiola M (2008) Neural substrates of early language acquisition. *Annu Rev Neurosci* 31:511–534
- Kuhl PK, Conboy BT, Padden D, Nelson T, Pruitt J (2005) Early speech perception and later language development: implications for the critical period. *Lang Learn Dev* 1:237–264
- Larsen ON, Goller F (1999) Role of syringeal vibrations in bird vocalizations. *Proc R Soc Lond B Biol Sci* 266:1609–1615
- Martins EP (2001) Tracker program for frame by frame analysis of video clips. Department of Biology, Indiana University, Bloomington. Distributed by the author at <http://www.indiana.edu/~martins/tracker.html>
- Moriyama K, Okanoya K (1996) Effect of beak movement in singing Bengalese finches. In: Abstracts of the third joint meeting, Acoustical Society of America and Acoustical Society of Japan, Honolulu, 2–6 December 1996, pp 129–130
- Nelson BS, Beckers GJL, Suthers RA (2005) Vocal tract filtering and sound radiation in a songbird. *J Exp Biol* 208:297–308
- Nowicki S (1987) Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature* 325:53–55
- Owerkovicz T, Farmer CG, Hicks JW, Brainerd EL (1999) Contribution of gular pumping to lung ventilation in monitor lizards. *Science* 284:1661–1663
- Podos J, Southall JA, Rossi-Santos MR (2004) Vocal mechanics in Darwin's finches: correlation of beak gape and song frequency. *J Exp Biol* 207:607–619
- Riede T, Beckers GJL, Blevins W, Suthers RA (2004) Inflation of the esophagus and vocal tract filtering in ring doves. *J Exp Biol* 207:4025–4036
- Riede T, Suthers RA, Fletcher NH, Blevins W (2006) Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc Natl Acad Sci USA* 103:5543–5548
- Suthers RA, Goller F (1997) Motor correlates of vocal diversity in songbirds. In: Nolan V Jr, Ketterson E, Thompson CF (eds) Current ornithology. Plenum Press, New York, pp 235–288
- Suthers RA, Zollinger SA (2008) From brain to song: the vocal organ and vocal tract. In: Zeigler P, Marler P (eds) The neuroscience of birdsong. Cambridge University Press, Cambridge (in press)
- Weary DM, Weisman RG (1992) Temporal relationships in white-throated sparrow song. *Condor* 94:1013–1016
- Westneat MW, Long J, John H, Hoese W, Nowicki S (1993) Kinematics of birdsong: functional correlation of cranial movements and acoustic features in sparrows. *J Exp Biol* 182:147–171
- Wild JM (1997) Neural pathways for the control of birdsong production. *J Neurobiol* 33:653–670