Elasticity and stress relaxation of rhesus monkey (*Macaca mulatta*) vocal folds

Tobias Riede

Department of Biology and National Center for Voice and Speech, University of Utah, Salt Lake City, UT 84112, USA

t.riede@utah.edu

SUMMARY

Fundamental frequency is an important perceptual parameter for acoustic communication in mammals. It is determined by vocal fold oscillation, which depends on the morphology and viscoelastic properties of the oscillating tissue. In this study, I tested if stress–strain and stress–relaxation behavior of rhesus monkey (*Macaca mulatta*) vocal folds allows the prediction of a species' natural fundamental frequency range across its entire vocal repertoire as well as of frequency contours within a single call type. In tensile tests, the load–strain and stress–relaxation behavior of rhesus monkey vocal folds and ventricular folds has been examined. Using the string model, predictions about the species' fundamental frequency range, individual variability, as well as the frequency contour of 'coo' calls were made. The low- and mid-frequency range (up to 2 kHz) of rhesus monkeys can be predicted relatively well with the string model. The discrepancy between predicted maximum fundamental frequency and what has been recorded in rhesus monkeys is currently ascribed to the difficulty in predicting the behavior of the *lamina propria* at very high strain. Histological sections of the vocal fold and different staining techniques identified collagen, elastin, hyaluronan and, surprisingly, fat cells as components of the *lamina propria*. The distribution of all four components is not uniform, suggesting that different aspects of the *lamina propria* are drawn into oscillation depending on vocal fold tension. A differentiated recruitment of tissue into oscillation could extend the frequency range specifically at the upper end of the frequency scale.

Key words: larynx, viscoelastic properties, bioacoustics, hyaluronan.

INTRODUCTION

The elongation of mammalian vocal folds causes the tissue to stiffen and therefore to vibrate at higher rates, directly affecting the fundamental frequency (f_0) of phonation (Hollien, 1960; Hirano et al., 1969). At strains larger than 20–30%, the stress response is highly nonlinear and demonstrates quantitative differences between individuals, sexes, as well as species (Haji, 1990; Min et al., 1995; Chan et al., 2007; Hunter and Titze, 2007; Zhang et al., 2009; Riede et al., 2010). When vocal folds are strained, some stress is released (Zhang et al., 2009) and would cause a drop in f_0 if not compensated, for example, by further stretching. In a first approximation, a vocal fold behaves like a string. Stiffening increases the oscillation rate and, over time, the oscillation rate decreases due to stress relaxation. The string (vocal fold) has to be tuned again. A string can be operated within a certain stress range. A minimum positive stress prevents the string from sagging, and a maximum stress indicates the point before the material gets irreversibly damaged. I tested the hypothesis that the string model allows the prediction of the f_0 range of a species' entire vocal repertoire as well as of the f_0 contour within a specific call. The rhesus monkey (*Macaca mulatta*) is an excellent model to study this hypothesis because its vocal repertoire is well known, and a vocalization such as the 'coo' call (Fig. 1) shows a very specific rising-and-falling f_0 contour that can only be achieved if laryngeal intrinsic muscles stiffen the vocal fold tissue to the correct stress level and simultaneously compensate for stress relaxation if necessary.

Most mammals produce their calls by vocal fold oscillations, which are caused by airflow forced through the larynx (e.g. Hast et al., 1974; Jürgens, 2002; Brown et al., 2003). The oscillation rate determines the f_0 of the sound that a con-specific can hear and interpret. f_0 is an important perceptual parameter for acoustic communication. It is controlled actively by the aerodynamic driving force (lung pressure), vocal fold positioning and nonlinear source–filter interactions, but,

most importantly, it is determined passively by viscoelastic properties of the oscillating tissue (Titze, 1988; Titze, 2008; Titze et al., 1997; Titze et al., 2008; Chan et al., 2009). The relationship between vocal fold viscoelastic properties and f_0 of the human voice has been established *via* the string model, in which string (or vocal fold) length, material stress and material density determine the rate of oscillation (e.g. Titze, 2000; Chan et al., 2009). In other species, it is unknown if and how the string model holds and how viscoelastic data relate to a species' f_0 range. In the present study, I investigated quality and magnitude of the vocal fold stress response in rhesus monkeys and made predictions about the species' expected f_0 range, specifically discussing the relevance of vocal fold viscoelastic properties for coo call production.

The coo call is a typical vocalization of all species in the genus *Macaca* (Green, 1975; Grimm, 1967; Itani, 1963; Lillehei and Snowdon, 1978; Hohmann and Herzog, 1985; Hauser, 1991; Hauser and Fowler, 1992; Owren et al., 1992; Owren and Rendall, 2003; Rendall et al., 1996). Coo calls are approximately 500 ms in duration. f_0 increases variably during the first part of the call. The f_0 contour varies at species, sexual and individual levels and is associated with certain context (Green, 1975; Hauser, 1991), and the laryngeal muscle activity demonstrates a f_0-dependent pattern (Hast et al., 1974; West and Larson, 1993).

The goal of the present study is to quantify the tensile stress response of male and female rhesus monkey vocal folds, including its time dependence. I have also collected histological data because the composition of the vocal fold, specifically of the lamina propria, is critical for its viscoelastic behavior. Additionally, I have investigated the ventricular folds. They mark the cranial edge of the entrance to the lateral laryngeal ventricles and therefore are very closely located to the vocal folds. In humans, the ventricular folds (or 'false folds') are involved in some types of phonation (e.g. Tsai
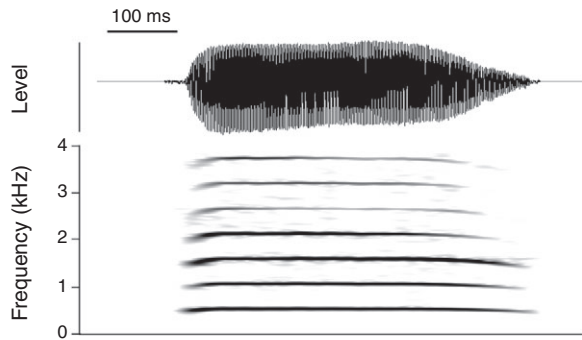
Fig. 1. Time signal and spectrogram of a rhesus macaque coo call (courtesy of Drew Rendall, University of Lethbridge). Note that the fundamental frequency in the coo call increases during the first 100 ms before it remains constant for a while and decreases towards the end. Stress relaxation in the monkey's vocal fold would likely cause the fundamental frequency to drop if no compensation for the tissue relaxation occurs.

et al., 2008). It is likely that they are also involved in nonhuman primate phonation, in particular in high-amplitude screaming (van den Berg, 1955; Agarwal et al., 2003; Alipour et al., 2007).

MATERIALS AND METHODS

Twelve rhesus monkey (*Macaca mulatta*, Zimmermann 1780) larynges (six males, six females) became available through the Wisconsin National Primate Research Center (WNPRC, Nonhuman Primate Tissue Distribution Program, Madison, WI, USA). Tissue was collected immediately after the animal was sacrificed. The tissue was quickly frozen in saline solution in liquid nitrogen and kept at –30°C until the experiment. Age, sex and body mass were available for all specimens (Table 1).

Histology

Six vocal and ventricular folds (three males, three females) were subjected to histology. The isolated tissue was stored in 10% neutral buffered formalin for 2 weeks before further processing. The tissue was then embedded in paraffin and 5 μm cross-sections were made.

Adjacent sections were exposed to one of the following stains: hematoxylin and eosin (H&E) for a general histological evaluation; elastica van Gieson stain (EVG) to identify elastic fibers; trichrome stain (TRI) to demonstrate collagen fibers; Alcian Blue stain (AB) (pH 2.5) to determine mucopolysaccharides and glycosaminoglycans. We also performed a hyaluronidase digestion (with bovine testicular hyaluronidase for 2 h at 37°C) in combination

Table 1. Summary of body mass (m_b) and vocal fold length (*L*) for all specimens tested

Animal	Age (years)	Sex	m_b (kg)	*L* (mm)
1	3	F	2.27	5.3
2	9	F	6.57	8.35
3	8	F	4.88	8.2
4	18	F	5.47	9.1
5	16	F	7.46	8.6
6	9.5	F	4.41	7.2
7	3.5	M	6.93	7.95
8	3.5	M	5.52	7.3
9	7	M	9.15	9.6
10	6	M	4.61	8.22
11	9	M	7.88	8.67
12	7	M	4.54	8.2

with a subsequent Alcian Blue stain to increase specificity for various acid mucosubstances in the Alcian Blue stain. Alcian Blue positivity is destroyed following prior incubation with hyaluronidase if hyaluronan is a major component of the mucosubstances. All stains were also performed on other tissues as positive control stains (artery for EVG; liver for TRI; small intestine for AB).

Stress–strain measurements

The force–elongation data were obtained by (a) a 1 Hz sinusoidal cyclic loading and (b) a stepwise loading procedure, both performed by an automated electromechanical system, recording force and distance.

Before dissection, the *in situ* vocal fold length was measured with an accuracy of 0.1 mm. One vocal fold and one ventricular fold from each of six males and six females were dissected using microsurgery instruments and were tested in their complete length. Vocal folds and ventricular folds remained attached to a small portion of the arytenoid cartilage dorsally and to the thyroid cartilage ventrally. The thyroarytenoid muscle was removed while the lamina propria remained intact. The epithelium was also carefully removed. Great care was taken to make sure that the remaining fibers were not damaged. One suture connected the arytenoid cartilage to the lever arm of the servo-control lever system, and a second connected the thyroid cartilage to a fixation point below the lever arm.

The tissue was vertically mounted in a water-surrounded chamber containing saline solution (Ringer solution) maintained at 38°C (Fig. 2). The exact length between cartilage and fixation point in the clamp was measured with a caliper (±0.1 mm accuracy).

The force–elongation data were obtained by 1 Hz sinusoidal stretch and release of the vocal fold by means of a dual-mode servo-control lever system (Model 305B; Aurora Scientific, Aurora, ON, Canada). Displacement of and force on the lever arm (resolution 1 μm and 0.3 mN) were recorded. Elongation was applied in a longitudinal direction (dorso-ventral), followed by a shortening to the original length. The present set of experiments was conducted with a system under displacement control. A controlled sinusoidal displacement was applied to the lever arm so that the vocal fold was lengthened and shortened (loading–unloading condition) 15 times at a frequency of 1 Hz. The force and elongation signals were then transmitted *via* a 16-bit analog-to-digital acquisition board (Windaq Model DI722; DATAQ Instruments, Akron, OH, USA) at 6 kHz sampling frequency to a PC.

After two minutes of rest, in a second set of experiments, stress relaxation under fixed strain was measured. For stress relaxation estimations, the tissue was strained to 50% (ventricular folds) and 60% (vocal folds), respectively, using 500 ms ramp time and a 20 s holding period.

A pre-strain of 20% was applied to each specimen. Pre-strain is the elongation (relative to total specimen length) imposed on the specimen prior to each lengthening–shortening test. Vocal fold length differed between *in situ* (intact larynx) and *ex situ* (vocal fold excised from the cartilage framework of the larynx), most likely due to suspension of the vocal folds in the laryngeal cartilage framework. A pre-strain of 20% compensated for length changes due to isolating the tissue and was based on measurements of vocal fold length before and after isolation from the cartilage framework. The tension rise during the 20% extension is negligible. The distance between insertion points at the thyroid cartilage and the arytenoid cartilage was 19.8±2.1% shorter (mean ± s.d., *N*=5) after isolating the vocal fold.

Mass measurements were affected by hydration (tissue bath, see above) of the isolated tissue. We corrected all mass measurements by 10%, accounting for mass gain, following earlier studies (Riede
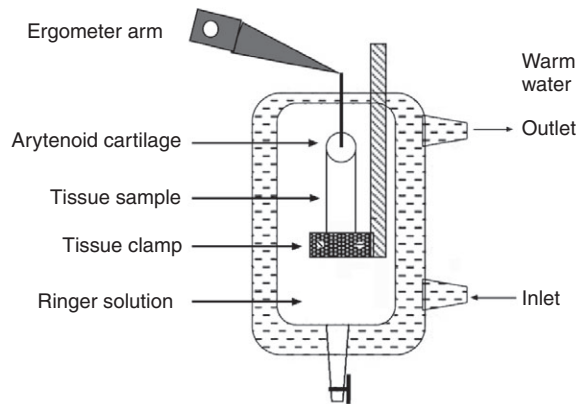
Fig. 2. The tissue sample was mounted by suturing the arytenoid cartilage to the ergometer arm. The other end of the tissue was a piece of thyroid cartilage. The thyroid cartilage was sutured to the metal frame that could also hold the clamp.

and Titze, 2008; Riede et al., 2010) as well as a measurement in one rhesus monkey vocal fold. We adapted density values from an earlier study (Min et al., 1995). Hunter et al. demonstrated that small variations in density have a negligible effect on stress calculations (Hunter et al., 2007).

Data analysis

True tensile strain (ε) was calculated as a logarithm of specimen length divided by its original mounting length:

$$\varepsilon = \ln\left(\frac{l}{l_0}\right), \tag{1}$$

where l_0 is the mounting length (including pre-strain) and l is the actual length of the specimen during stretching. Note that when strain is given as percentage in this paper, the term on the right side of Eqn 1 is multiplied by 100.

True tensile stress (σ) is defined as the ratio between force (F; in N) and cross-sectional area (A; in m^2) of the specimen. Assuming tissue incompressibility, uniform specimen cross-sectional area with roughly cylindrical geometry, and tissue isotropy allow for calculating the average cross-sectional area. The mean cross-sectional area ($\overline{A}$) is then:

$$\overline{A} = \frac{m}{\rho(l_0 + \Delta l)}, \tag{2}$$

where m is the specimen mass and ρ is the tissue density. This equation considers the strain-dependent cross-sectional area A, which decreases as the specimen is elongated and increases as the specimen returns to its initial mounting length. Usually, 'engineering' or 'nominal stress' is used (where the cross-sectional area, A, is assumed constant). However, when the changes in cross-sectional area are significant, then stress must be calculated using the strain-dependent cross-sectional area called 'true stress'. Considering the large strains applied to the tissue, the use of true stress is critical. With the varying cross-sectional area, the tensile stress (in Pa) can be calculated as:

$$\sigma = \frac{F\rho(l_0 + \Delta l)}{m}, \tag{3}$$

where F is the applied force.

The ratio of stress and strain tells us how much the material stretches for a given load. It is called variously 'stiffness', 'elastic modulus' or 'Young's modulus of elasticity'.

Cyclic loading

The overall stress–strain response of vocal fold tissue has been differentiated into a linear low-strain and a nonlinear high-strain region. The low-strain region is modeled with a linear function:

$$\sigma = a\varepsilon + b \quad \text{for} \quad \varepsilon < \varepsilon_1, \tag{4}$$

where a is the slope of the curve, and b is the y-axis intercept. The high-strain region is best approximated with an exponential equation:

$$\sigma = Ae^{B\varepsilon} \quad \text{for} \quad \varepsilon \geq \varepsilon_1, \tag{5}$$

where A and B are constants to be determined empirically.

The amplitude of the cyclic strain application remains constant during 15 cycles, and the amplitude of the resulting stress response decreases (Fig. 3A). The combined stress–strain response results in a 'banana-shaped' curve (Fig. 3B), with a loading part (upper curve) higher than the unloading part (lower curve). The constants in Eqns 4 and 5, as well as the upper limit of the linear strain region ('linear strain limit', ε_1), are derived by fitting a linear and an exponential regression line, respectively, to the empirical data, while maximizing the sum of both regression coefficients. The maximization process is performed by conducting a step-wise movement of the linear strain limit and performing a linear and exponential regression, respectively, on the two data sets at each step. The two resulting regression coefficients are added at each step [see Riede and Titze (Riede and Titze, 2008) for more detail]. The maximum of the sum of both regression coefficients is considered the linear strain limit, ε_1, in Eqns 4–6.

The range of Young's modulus was described by the first derivative of the respective stress–strain function:

$$\begin{aligned} E &= a & \text{for} \quad \varepsilon < \varepsilon_1, \\ E &= \frac{d\sigma}{d\varepsilon} = ABe^{B\varepsilon} & \text{for} \quad \varepsilon \geq \varepsilon_1, \end{aligned} \tag{6}$$

where A, B and a are constants.

In a 1 Hz cyclic tensile test, an energy loss in the unloading phase relative to the loading phase was observed in all specimens (i.e. difference between upper and lower curve in Fig. 3B). This energy loss, or hysteresis, was estimated as the difference between the area under the curves of the loading and the unloading phases of the 1 Hz-sinusoidal stress–strain response. Hysteresis is frequency dependent because it involves the strain rate with respect to time. We provide here only one estimate for a 1 Hz loading–unloading regime. The stress–strain responses (loading phase and unloading phase separately) were fitted with an exponential curve according to Eqn 5. The area under the curve between ε_1 and maximum applied strain was estimated by integrating both exponential equations. The difference in the area under the curve between loading and unloading stress–strain response was considered hysteresis and expressed as a percentage.

Stress relaxation

Vocal and ventricular fold tissue is stretched at a predetermined rate (here 0.5 Hz) to the desired strain (50% for ventricular folds and 60% for vocal folds). The tissue is maintained at the respective strain for 20 s, thus producing a stress relaxation curve (Fig. 4). The curve was modeled with an exponential decay function between the time
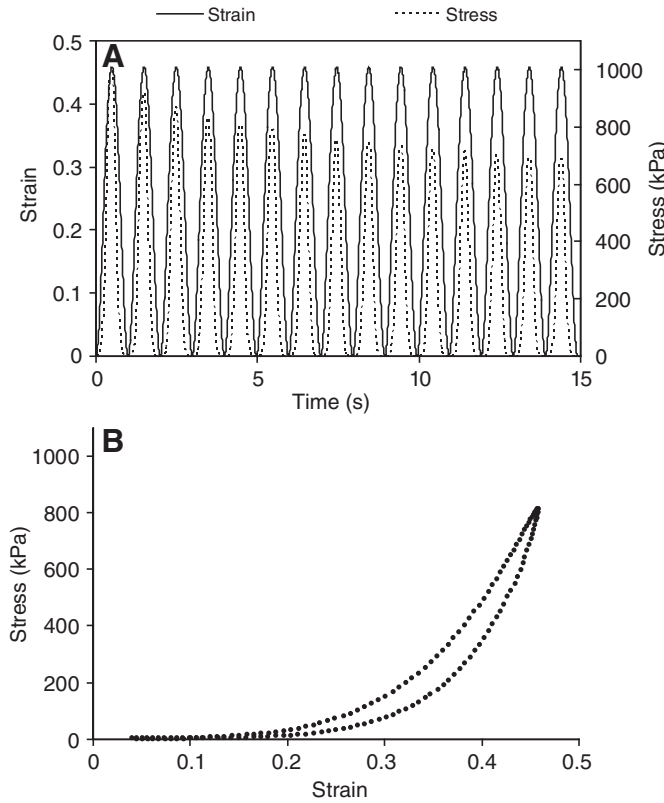
Fig. 3. Stress–strain response in time from a 1 Hz sinusoidal elongation of lamina propria. (A) Note that the amplitude of strain remains constant while stress decreases over time. The decrease in stress is a result of stress relaxation. (B) Stress–strain relationship for a single cycle from the same data set. The upper part of the 'banana-shaped' curve is the loading phase (stretching). The lower part is the unloading phase (return to the original mounting length). The difference between both curves is due to hysteresis of the tissue, i.e. lower stress in the tissue during the unloading phase. The low strain regions were fitted with a linear regression line, while the high-strain regions of both curves were modeled with exponential functions. In the loading curve, the limit of the linear region is the 'linear strain limit' (ε_1), here approximately 0.16.

point when the maximum stress was achieved and one second thereafter:

$$\sigma(t) = (\sigma_0 - P) \times e^{-Kt} + P, \qquad (7)$$

where P is the estimated plateau, K is the rate constant and σ_0 is the peak stress at time point zero (t_0). Stress half-life is computed as $\ln(2)/K$.

Viscoelastic properties were tested for differences between males and females as well as between vocal folds and ventricular folds using t-tests.

Fundamental frequency predictions

The anatomical arrangement of the vocal fold inside the larynx invites comparison of vocal fold oscillation with that of a string. The vocal fold is positioned between the thyroid cartilage and the arytenoid cartilage, and dorso-ventral elongation increases its tension. The airflow provides the energy source to set vocal folds into vibration. The rate of vibration determines f_0 and depends on vocal fold length and viscoelastic properties (van den Berg, 1958; Titze, 2006).
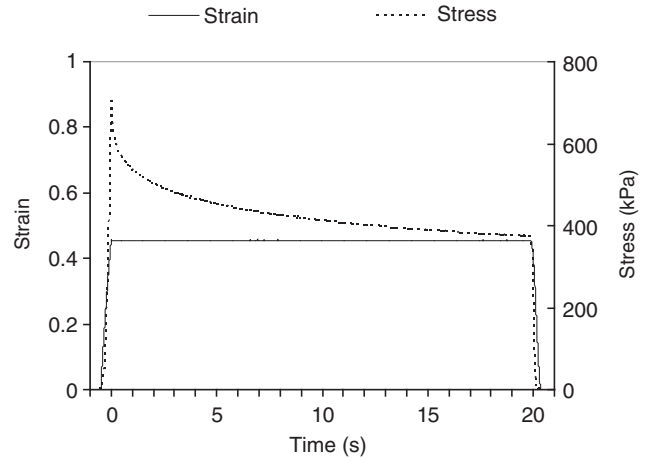

Fig. 4. Load-strain and stress relaxation curve. The vocal fold tissue is stretched at a predetermined rate (here to 50% and 60%, respectively, within 500 ms) to the desired strain, at which it is held for 20 s.

According to the string model, f_0 is determined by:

$$f_0 = \frac{1}{2L} \sqrt{\frac{\sigma}{\rho}}, \qquad (8)$$

where L is the string (or vocal fold) length, σ is the tissue stress (force per unit area), and ρ is the tissue density ($1.02\,\mathrm{g\,cm^{-3}}$). The stress response during longitudinal straining was tested here.

The stress–strain data collected here, together with the string model, were used to make inferences about f_0. The cyclic loading data were used to estimate a f_0 range. The f_0 curves for all individuals and the mean male and female vocal fold were calculated.

The peak stress data from each of the 15 cycles in the cyclic loading experiment were used to estimate f_0 relaxation due to repeated straining.

The stress relaxation data were used to estimate the acoustic effect of tissue relaxation on f_0. f_0 curves for all individuals were calculated implementing the stress relaxation curve in Eqn 8. Individual vocal fold lengths were measured before dissections (Table 1).

RESULTS
Histology

Vocal folds consist of thyroarytenoid muscle, lamina propria and epithelium. Ventricular folds consist of lamina propria and epithelium. The extracellular matrix of vocal and ventricular fold consists of collagen, elastin fibrillar proteins and hyaluronan (Fig. 5). The removal of hyaluronan by hyaluronidase digestion, with subsequent Alcian Blue staining, indicated that much of the positive stain in the vocal fold (Fig. 5E) and the ventricular fold (Fig. 5I) can be attributed to hyaluronan (Fig. 5F,J). Laryngeal glands, located in the ventricular fold, remained intensely positively stained after hyaluronidase treatment. Although the glucosaminoglycans of vocal and ventricular fold were mainly made of hyaluronan, which was digested away by hyaluronidase, it was a composition of various mucopolysaccharides and glycosaminoglycans in the laryngeal gland. This pattern was found in all six specimens. In all six investigated specimens, fat was found in the deep portion of the lamina propria of the vocal fold and throughout that of the ventricular fold (Fig. 5E,I).
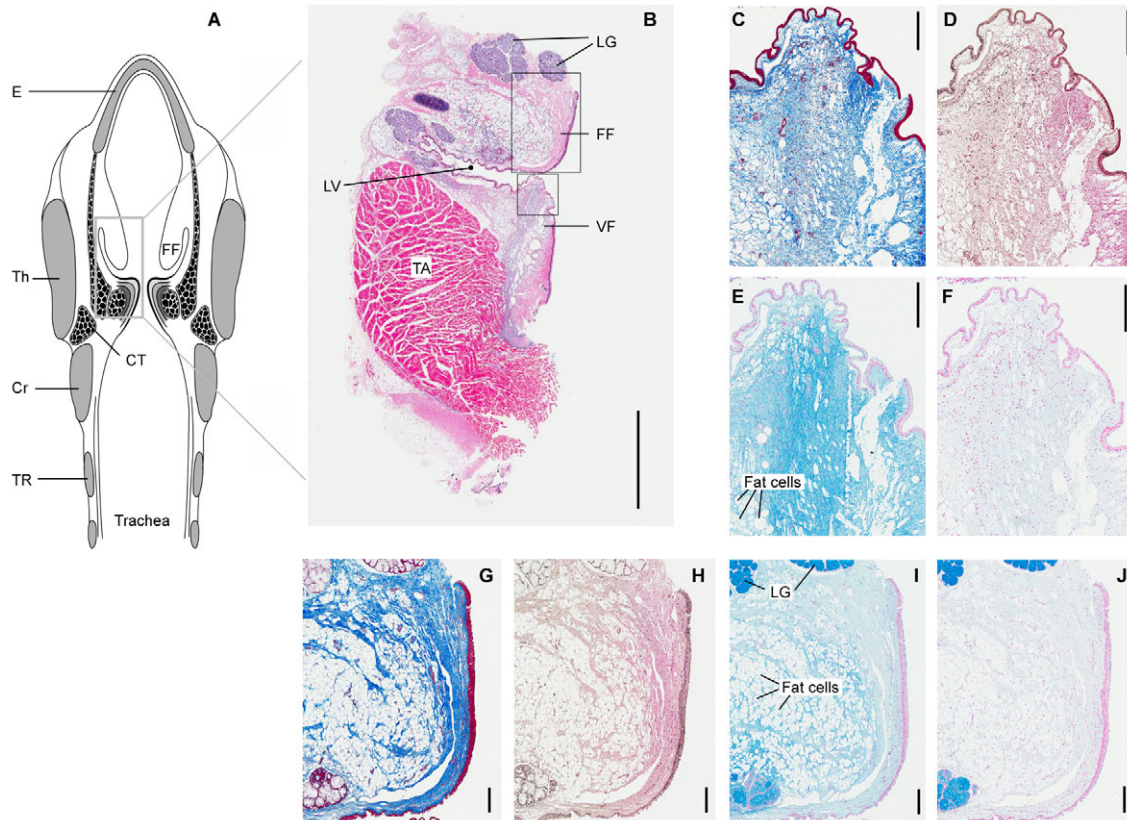
Fig. 5. Histological sections of the mid-membranous part of the vocal fold of a 6-year-old male rhesus monkey. (A) Schematic horizontal section of a larynx, indicating where the histological sections were taken from (gray box). (B) Hematoxylin and eosin stain of vocal fold (VF) and ventricular fold (FF); the two boxes indicate the position of vocal fold (C–F) and ventricular fold sections (G–J). C and G are trichrome stains, indicating collagen fibers in blue; D and H are elastica van Gieson stains, indicating elastic fibers in black stain; E and I are Alcian Blue stains; F and J are Alcian Blue stains after hyaluronidase digestion. Scale bar in B, 2 mm; scale bars in C–J, 200 μm. TA, thyroarytenoid muscle; CT, cricothyroid muscle; E, epiglottis; Th, thyroid cartilage; Cr, cricoid cartilage; TR, tracheal ring; LV, laryngeal ventricle; LG, laryngeal gland.

Dorso-ventral longitudinal stress–strain relationship

Rhesus monkey vocal folds and ventricular folds showed the typical linear stress–strain response in the low-strain region and a nonlinear relationship in the high-strain region (Fig. 3B). Linear and nonlinear models (Eqns 4 and 5) reached regression coefficients of 0.97 and higher (Table 2).

The linear strain limit (ε_1) was not significantly different between males and females, for either vocal folds or ventricular folds (Table 2). The slope of the linear model (constant a), the y-intercept (constant b) (Eqn 4) and the A constant in the exponential model (Eqn 5) were also not significantly different between males and females (Table 2). The B constant of the exponential model (Eqn 5) was significantly larger in males (Table 2), leading to a higher modulus at strains larger than ε_1, meaning the vocal folds from six male rhesus monkeys were, on average, stiffer than that of the six females at strains above the linear limit.

The linear and exponential models of the stress–strain response of vocal folds and ventricular folds were compared pair-wise in order to test for differences between those two structures within individuals. Only the b constant of the exponential model in females showed significant differences (Table 3). In females, the ventricular fold reaches systematically higher moduli at strains larger than the linear strain limit.

The between-individual coefficients of variations (CV) were calculated (Table 2). For vocal folds, the CV values are tentatively larger in females than in males (t-test; $t=-2.18$, $P=0.081$); for ventricular folds, the CV values are not significantly different (t-test; $t=0.07$, $P=0.94$).

Hysteresis ranged from 27 to 36% for male vocal folds, from 27 to 40% for female vocal folds, from 27 to 40% for male ventricular folds and from 27 to 44% for female ventricular folds (Table 2). Differences between male and female tissue (Table 2), as well as between vocal and ventricular fold within each sex (Table 3), were not significant.

Stress relaxation

The stress relaxation in the stepwise procedure (Fig. 4) within the first second after reaching peak stress ranged from 22 to 35% for male vocal folds, from 17 to 46% for female vocal folds, from 16 to 29% for male ventricular folds and from 15 to 29% for female ventricular folds (Table 4). The stress half-life ranged from 147 to 263 ms for male vocal folds, from 94 to 265 ms for female vocal folds, from 193 to 236 ms for male ventricular folds and from 188 to 236 ms for female ventricular folds (Table 4). Differences between male and female tissue were not significant (Table 4).

Fundamental frequency predicted by biomechanical model

Data were implemented in a combination of the linear and exponential models (Eqns 4 and 5) using individual stress responses and vocal fold length (Table 1). For the average stress–strain curve, between the original vocal fold length (which was estimated at 7.8 mm for female and 8.3 mm for male vocal folds, according to

Table 2. Parameters of the linear (Eqn 4) and the exponential (Eqn 5) model for curve-fitting the empirical stress–strain response of the vocal fold

		Males	CV (%)	Females	CV (%)	t-test, males vs females
Vocal folds	A	52.0±21.5	41.3	54.6±31.1	56.9	t=0.16; P=0.87
	B	0.4±0.2	50.0	0.4±0.5	125	t=−0.09; P=0.92
	R^2	0.97±0.009		0.97±0.001		
	ε_1	13.3±3.5	25.9	13.4±5.3	39.5	t=0.03; P=0.96
	a	1.06±0.7	66.0	2.5±2.1	84.0	t=1.53; P=0.15
	b	16.2±1.3	8.0	12.9±2.4	18.6	**t=−2.88; P=0.016**
	R^2	0.99±0.003		0.99±0.009		
	H	31.5±3.1	9.8	34.1±5.2	15.2	t=1.06; P=0.31
Ventricular folds	A	57.7±59.4	102.9	28.1±13.4	47.7	t=−1.2; P=0.26
	B	0.2±0.1	50.0	0.3±0.4	133.3	t=0.56; P=0.58
	R^2	0.96±0.01		0.97±0.009		
	ε_1	10.5±3.1	29.5	13.3±3.3	24.8	t=1.48; P=0.16
	a	3.1±3.9	125.8	0.9±0.9	100.0	t=−1.33; P=0.21
	b	14.7±3.7	25.2	16.4±2.5	15.2	t=0.89; P=0.39
	R^2	0.97±0.01		0.98±0.01		
	H	35.1±5.0	14.2	35.8±6.5	18.1	t=0.21; P=0.83

All means and standard deviations are based on six samples. R^2, regression coefficient; ε_1, linear strain limit; H, hysteresis. Significant results are highlighted in bold.

data in Table 1) and 50% elongation, the string model suggests a maximum f_0 range between ~80 Hz and 1.6 kHz (Fig. 6).

The relaxation of peak stress in 15 cycles can lead to a 10 to 30% drop in fundamental frequency (Fig. 7).

The stress relaxation within 1 s after reaching peak stress in a stepwise procedure to 50% strain can lead to a 10 to 30% drop in fundamental frequency (Fig. 8).

DISCUSSION

Rhesus monkey vocal folds and ventricular folds respond almost linearly with stress at strains up to ~15%, i.e. elongation leads to a proportional increase in stiffness and in f_0. Beyond ~15% strain, the stress response is nonlinear. Small length changes result in large stiffness changes and therefore would result in large f_0 changes. The stress response is different in male and female vocal folds. This parallels similar sex-specific differences in human (Chan et al., 2007) and Rocky Mountain elk (Riede and Titze, 2008) vocal folds. In all three species, it is the male vocal fold that is somewhat stiffer, suggesting that testosterone mediates higher tissue stiffness in vocal folds. Ventricular folds of rhesus monkeys demonstrate no sex differences, unlike human ventricular folds (Chan et al., 2006). This

Table 3. Test results of a pair-wise comparison of the linear (constants A and B) and exponential (constants a and b) model parameters between vocal folds and ventricular folds

		t	P
Males	A	−0.23	0.82
	B	1.86	0.12
	ε_1	1.33	0.24
	a	−1.26	0.26
	b	0.73	0.49
	H	−1.39	0.22
Females	A	1.89	0.11
	B	0.29	0.77
	ε_1	0.09	0.92
	a	1.75	0.13
	b	**−3.35**	**0.02**
	H	−0.59	0.57

ε_1, linear strain limit; H, hysteresis. Significant results are highlighted in bold.

difference between human and rhesus monkey tissue could be based on their different morphology. The extracellular matrix of the former consists of collagen, elastic fibers and hyaluronan in which gland tissue is embedded, and the latter contains additionally large amounts of fat cells (Fig. 5). Fat in vocal folds has been interpreted as an adaptation to produce low f_0 by enlarging the oscillating portion of the vocal fold (Hast, 1989). It is also injected as graft to substitute tissue in a damaged human vocal fold (Mikaelian et al., 1991). Its function in rhesus monkeys is not clear.

How do tissue properties affect vocalization; for example, the coo call production? Its f_0 contour is individual specific (Hauser, 1991; Owren and Rendall, 2003). Sources of f_0 individual specificity include vocal fold morphology (size, shape and viscoelastic properties) and central control of laryngeal and respiratory muscles. The lamina propria stiffness shows individual differences, particularly at high strains, which could result in a f_0 difference on the magnitude of 2 at 50% strain (Fig. 6). Human vocal folds also show a large inter-subject variability in their stress response in tensile tests (Zhang et al., 2009). However, certain call types, such as the coo call, are produced within a narrow f_0 range. In order to achieve an across-individual similarity in the f_0 range, vocal folds must be operated taking their viscoelastic properties into account. Data by Nishizawa et al. suggest that, in humans, length changes of the vocal fold that accompany the change of f_0 are highly individual-specific (Nishizawa et al., 1988). These authors report that, in human subjects, vocal fold length may vary as little as 21% and as much as 111% over an individual's entire f_0 range (Nishizawa et al., 1988). Their data also suggest that the length change required for producing a target f_0 change differs significantly between subjects.

The knowledge about the viscoelastic properties of vocal folds allows us to make predictions about not only constraints but also the costs of sound production. Coo calls are produced within a relatively narrow f_0 range. If the species' predicted total f_0 range is between 80 and about 2000 Hz (Fig. 6), then coo calls with a f_0 ranging between 200 and 600 Hz (Hauser, 1991; Owren and Rendall, 2003) must be produced with vocal folds strained to about 15–35% (Fig. 6). This is close to the linear strain limit (Table 2), suggesting that coo call production requires only a small to medium effort by intrinsic laryngeal muscles.

Table 4. Parameters of exponential decay model (Eqn 7) for curve-fitting the empirical stress relaxation curve of the vocal fold, as well as stress half life and relative energy loss after one second

		Males	CV (%)	Females	CV (%)	t-test, males vs. females
Vocal folds	σ_0	1372±563		784±330		t=−2.04; P=0.07
	P	1077±459		568±232		t=−2.2; P=0.053
	R^2	0.98±0.005		0.98±0.004		
	K	3.39±0.74		3.97±1.94		t=0.07; P=0.51
	HT	0.211±0.04		0.201±0.07		t=−0.31; P=0.76
	E	25.5±5.0		29.6±10.5		t=−0.85; P=0.41
Ventricular folds	σ_0	861±319		799±326		t=−0.29; P=0.77
	P	708±288		631±248		t=−0.43; P=0.67
	R^2	0.96±0.02		0.96±0.01		
	K	3.32±0.22		3.25±0.34		t=−0.24; P=0.67
	HT	0.209±0.01		0.214±0.02		t=0.5; P=0.62
	E	22.2±4.4		24.0±6.3		t=0.53; P=0.60

σ_0, peak stress (in kPa); P, plateau stress (in kPa); K, rate constant (in kPa s^{-1}); R^2, regression coefficient; HT, half-life (in s); E, relative energy loss (in %). Each vocal fold was stretched to 60% strain, and each ventricular fold was stretched to 50% strain. Each mean and s.d. are based on six (males) or five (females) samples.

How well does the predicted and observed f_0 range overlap? The lowest f_0 of 80–100 Hz corresponds to what has been reported in the literature (Green, 1975; Grimm, 1967; Lillehei and Snowdon, 1978; Hohmann and Herzog, 1985; Hauser, 1991; Hauser and Fowler, 1992; Owren et al., 1992; Owren and Rendall, 2003; Patel and Owren, 2007). The estimated maximum f_0 of ~2 kHz fits many harmonic call types reported for rhesus monkeys, such as coo calls (Rowell and Hinde, 1962; Hauser, 1991; Owren and Rendall, 2003) or the girney call (Rowell and Hinde, 1962; Hauser and Fowler, 1992). But the highest ever reported f_0 is close to 8 kHz for screams and squeaks (e.g. Rowell and Hinde, 1962; Gouzoules et al., 1984). These are high amplitude, often containing noisy and harmonic components. At this stage, it is hard to explain the discrepancy between expected and reported maximum f_0. Possible explanations include the involvement of additional tissue into oscillation during screaming, like ventricular folds, for example, in humans (Ufema and Montequin, 2001), which might be able to oscillate at higher frequencies or are involved in nonlinear interactions (Titze et al., 2008). A very important alternative explanation is the reduction of the amount of oscillating tissue at highest frequencies. Among mammals, the upper oscillation rate limit is difficult to predict because the vocal fold is a layered structure (Hirano, 1975), and each layer has its own specific viscoelastic properties (Zhang et al., 2009). Therefore, it is hard to predict how much tissue will be drawn into oscillation and what effective mass and effective stress will be present at these highest strains. Here, the stress response of the complete lamina propria was tested, but during phonation epithelium and variable amounts of lamina propria are drawn into oscillation (Hirano, 1975). Rhesus monkey vocal folds consist of thyroarytenoid muscle, lamina propria and epithelium (Fig. 5). The lamina propria consists of elastin, collagen and hyaluronan (Fig. 5) and, unlike human (e.g. Hirano 1974; Hahn et al., 2006a; Hahn et al., 2006b) and many other mammal vocal folds (e.g. Kurita et al., 1983; Riede et al., 2010), fat cells (Fig. 5E). All four components are not equally distributed and are likely contributing to differentiated viscoelastic properties throughout the lamina propria, giving rise to functionally different layers in the lamina propria.

Rhesus monkey vocal fold tissue relaxes over time at constant strain (Fig. 4). The short- and long-term tissue time-dependence would affect f_0 by lowering the tensile stress, accounting for a significant portion of f_0 drop. In rhesus monkeys, this could lower f_0 by 20–30% within a second (Fig. 8). However, a typical coo call shows a f_0 rise from the beginning up to a certain time point within

the call (Fig. 1). The modulation during the rise and the position of the peak f_0 are reportedly associated with context (Hauser, 1991; Owren and Rendall, 2003). Muscle activity must counter the tissue's tendency to relax in order to achieve a f_0 rise. Electromyography of the intrinsic laryngeal muscles in rhesus monkeys responsible for vocal fold elongation are typically activated about 100–200 ms before the coo call onset (Hast et al., 1974; West and Larson, 1993; Jürgens, 2002). The prephonatory muscle activity is of the same temporal magnitude as the stress half-life of the lamina propria (Table 4). It is tempting to conclude that prephonatory muscle
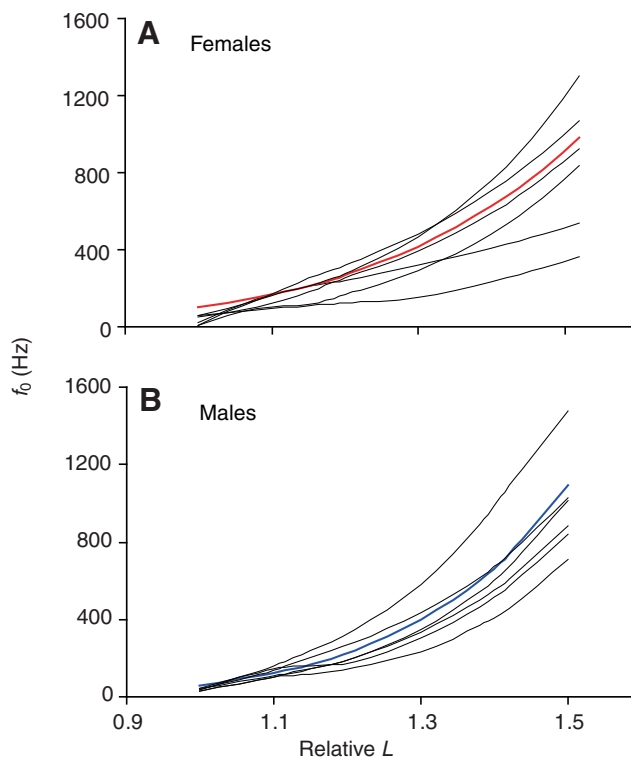

Fig. 6. Fundamental frequency (f_0) predicted by the string model (Eqn 8). The linear and nonlinear stress responses for female (A) and male (B) vocal folds were implemented in Eqn 8 for strains between 0 and ~50%. The red and blue lines, respectively, indicate the mean relationship between vocal fold length (L) and f_0.
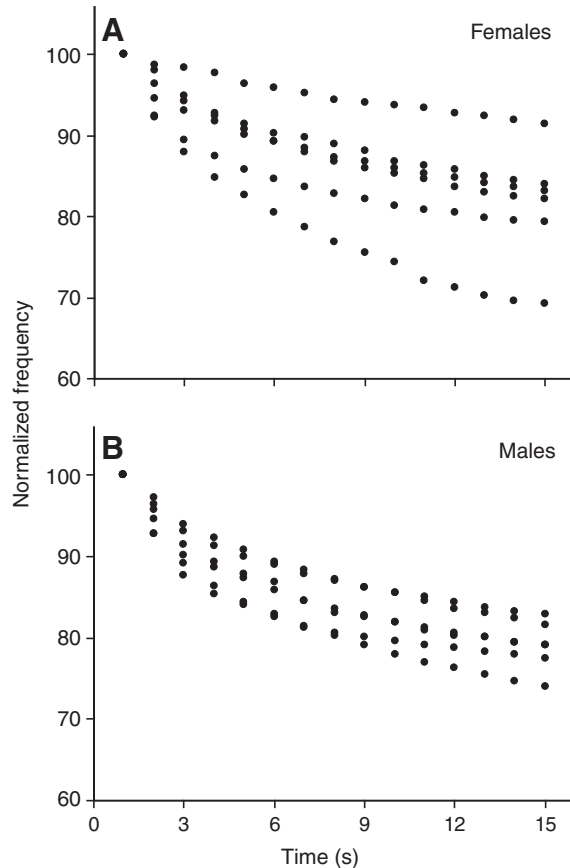
Fig. 7. Fundamental frequency at peak stress in 15 loading cycles, predicted by the string model (Eqn 8) in six females (A) and six males (B).
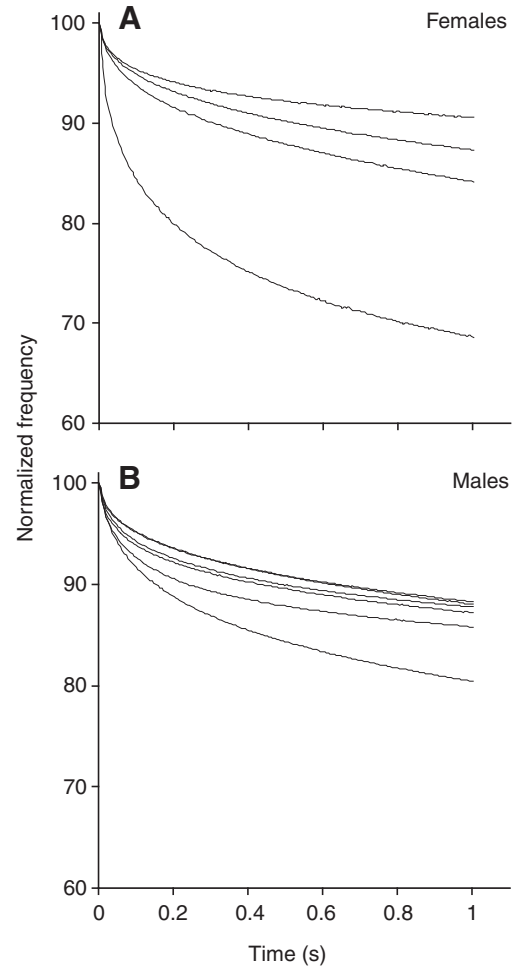

Fig. 8. Fundamental frequency development after stretching the vocal fold to 40% strain (at time point zero), predicted by the string model (Eqn 8) in four females (A) and six males (B). The mean stress relaxation curve is indicated in red and blue, respectively.

activity is not only responsible for positioning vocal folds but also serves to precondition the vocal fold tissue and thereby account for the initial very dramatic tissue relaxation after reaching the peak stress.

Rhesus monkey vocal fold tissue also relaxes from loading cycle to loading cycle (Fig. 3A and Fig. 7). After 15 cycles, the simulated f_0 has dropped by 10–20%, which is comparable to data found in humans (Chan et al., 2009). Long-term f_0 decrease in human speech is called f_0 declination. f_0 declination refers to an observation in many languages in which f_0 declines in a phrase or utterance (e.g. Bolinger, 1978). Hauser and Fowler (Hauser and Fowler, 1992) found f_0 declination to occur in call sequences of rhesus monkeys and vervet monkeys (*Cercopithecus aetiops*). One suggestion for a functional relevance of this phenomenon is that f_0 declination serves as a temporal cue for an indication of the end of an utterance (Breckenridge, 1977). A listener could use f_0 development as an indicator for boundaries. Hauser and Fowler suggested that constraints of the sound production apparatus might account for the occurrence of this phenomenon in humans and nonhuman primates (Hauser and Fowler, 1992). Since f_0 and viscoelastic properties of the vocal fold are tightly connected, stress relaxation in vocal folds could be an important factor contributing to f_0 declination in humans (Chan et al., 2009) as well as in nonhuman primates (present study). The fact that viscoelastic properties are individual-specific could also explain why results about f_0 declination can be mixed (for a review, see Hauser and Fowler, 1992). Of course, f_0 regulation is more complex because it is also affected by other factors such as aerodynamic driving pressure (lung pressure). For example, 't Hart et al. demonstrated that lung pressure can decrease over the course of an utterance and could therefore contribute to f_0 declination ('t Hart et al., 1990).

Sound production is complex, involving the coordination of different motor patterns and physical properties. Among the passive physical properties, the vocal fold stress response and its time dependence represent an important constraint of laryngeal sound production in human and nonhuman primates affecting the short- and long-term f_0 contour of vocalization. It therefore seems to be a feature of universal relevance in many primates and is not unique to the human language faculty.

LIST OF SYMBOLS AND ABBREVIATIONS

A	cross-sectional area
$\overline{A}$	mean cross-sectional area
AB	Alcian Blue stain
CV	coefficient of variation
EVG	elastica van Gieson stain
F	force
f_0	fundamental frequency
H&E	hematoxylin and eosin stains
L	string (or vocal fold) length

l	specimen length
l_0	original mounting length
m_b	body mass
t	time
TRI	trichrome stain
ε	tensile strain
ε_1	linear strain limit
σ	tensile stress
σ_0	peak stress at time point zero

REFERENCES

Agarwal, M., Scherer, R. C. and Hollien, H. (2003). The false vocal folds: shape and size in frontal view during phonation based on laminagraphic tracings *J. Voice* **17**, 97-113.

Alipour, F., Jaiswal, S. and Finnegan, E. (2007). Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models. *Ann. Otol. Rhinol. Laryngol.* **116**, 135-144.

Bolinger, D. (1978). Intonation across languages. In *Universals of Human Language: Phonology* (ed. J. Grennberg), pp. 471-524. Stanford: Stanford University Press.

Breckenridge, J. (1977). *Declinationas a Phonological Process.* Technical Memo. Murray Hill, New Jersey: Bell Lab.

Brown, C., Alipour, F., Berry, D. A. and Montequin, D. (2003). Laryngeal biomechanics and vocal communication in the squirrel monkey (*Saimiri boliviensis*). *J. Acoust. Soc. Am.* **113**, 2114-2126.

Chan, R. W., Fu, M. and Tirunagari, N. (2006). Elasticity of the human false vocal fold. *Ann. Otol. Rhinol. Laryngol.* **115**, 370-381.

Chan, R. W., Fu, M., Young, L. and Tirunagari, N. (2007). Relative contributions of collagen and elastin to elasticity of the vocal fold under tension. *Ann. Biomed. Eng.* **35**, 1471-1483.

Chan, R. W., Siegmund, T. and Zhang, K. (2009). Biomechanics of fundamental frequency regulation: constitutive modeling of the vocal fold lamina propria. *Logopedics Phoniatrics Vocology* **34**, 181-189.

Gouzoules, S., Gouzoules, H. and Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) agonistic screams: representational signalling in the recruitment of agonistic aid. *Anim. Behav.* **32**, 182-193.

Green, S. (1975). Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): a field study. In *Primate Behavior* (ed. L. A. Rosenblum), pp. 1-102. New York: Academic Press.

Grimm, R. J. (1967). The vocal communication of pigtail macaques (*Macaca nemestrina*). *J. Zool.* **152**, 361-373.

Hahn, M. S., Kobler, J. B., Zeitels, S. M. and Langer, R. (2006a). Quantitative and comparative studies of the vocal fold extracellular matrix-I: Elastic fibers and hyaluronic acid. *Ann. Otol. Rhinol. Laryngol.* **115**, 156-164.

Hahn, M. S., Kobler, J. B., Zeitels, S. M. and Langer, R. (2006b). Quantitative and comparative studies of the vocal fold extracellular matrix II: Collagen. *Ann. Otol. Rhinol. Laryngol.* **115**, 225-232.

Haji, T. (1990). Mechanical properties of the vocal fold. *Practica Otologica Kyoto* **83**, 793-808.

Hast, M. H. (1989). The larynx of roaring and non-roaring cats. *J. Anatomy* **163**, 117-121.

Hast, M. H., Fischer, J. M., Wetzel, A. B. and Thompson, V. E. (1974). Cortical motor representation of the laryngeal muscles in *Macaca mulatta. Brain Res.* **73**, 229-240.

Hauser, M. D. (1991). Sources of acoustic variation in rhesus macaque (*Macaca mulatta*) vocalizations. *Ethology* **89**, 29-46.

Hauser, M. D. and Fowler, C. A. (1992). Fundamental frequency declination is not unique to human speech: evidence from nonhuman primates. *J. Acoust. Soc. Am.* **91**, 363-369.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr. Basel* **26**, 89-94.

Hirano, M. (1975). Phonosurgery and clinical investigations. *Otologia* (*Fukuoka*) **21**, 239-440.

Hirano, M., Ohala, J. and Vennard, W. (1969). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *J. Speech Hear. Res.* **12**, 616-628.

Hohmann, G. and Herzog, M. O. (1985). Vocal communication in liontailed macaques. *Folia primatol.* **45**, 148-178.

Hollien, H. (1960). Vocal pitch variation related to changes in vocal fold length. *J. Speech Hear. Res.* **3**, 150-156.

Hunter, E. J. and Titze, I. R. (2007). Refinements in modeling the passive properties of laryngeal soft tissue. *J. Appl. Physiol.* **103**, 206-219.

Hunter, E. J., Alipour, F. and Titze, I. R. (2007). Sensitivity of elastic properties to measurement uncertainties in laryngeal muscles with implications for voice fundamental frequency prediction. *J. Voice* **21**, 641-650.

Itani, J. (1963). Vocal communication of the wild Japanese monkeys. *Primates* **4**, 11-66.

Jürgens U. (2002). Neural pathways underlying vocal control. *Neurosci. Behav. Rev.* **26**, 235-258.

Kurita, S., Nagata, K. and Hirano, M. (1983). A comparative study of the layer structure of the vocal fold. In *Vocal Fold Physiology, Contemporary Research and Clinical Issues* (ed. D. M. Bless and J. H. Abbs), pp. 3-21. San Diego, USA: College Hill Press.

Lillehei, R. A. and Snowdon, C. T. (1978). Individual and situational differences in the vocalizations of young stumptail macaques. *Behaviour* **65**, 270-281.

Mikaelian, D. O., Lowry, L. D. and Sataloff, R. T. (1991). Lipoinjection for unilateral vocal cord paralysis. *Laryngoscope* **101**, 465-468.

Min, Y. B., Titze, I. R. and Alipour-Haghighi, F. (1995). Stress-strain response of the human vocal ligament. *Ann. Otol. Rhinol. Laryngol.* **104**, 563-569.

Nishizawa, N., Sawashima, M. and Yonemoto, K. (1988). Vocal fold length in vocal pitch change. In *Vocal Physiology: Voice Production, Mechanisms and Functions.* (ed. O. Fujimura), pp. 75-82. New York: Raven Press.

Owren, M. J. and Rendall, D. (2003). Salience of caller identity in rhesus monkey (*Macaca mulatta*) coos and screams: perceptual experiments with human (*Homo sapiens*) listeners. *J. Comp. Psychol.* **117**, 380-390.

Owren, M. J., Dieter, J. A., Seyfarth, R. M. and Cheney, D. L. (1992). 'Food' calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species. *Behaviour* **120**, 218-231.

Patel, E. R. and Owren, M. J. (2007). Acoustics and behavioral contexts of "gecker" vocalizations in young rhesus macaques (*Macaca mulatta*). *J. Acoust. Soc. Am.* **121**, 575-585.

Rendall, D., Rodman, P. S. and Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Anim. Behav.* **51**, 1007-1015.

Riede, T. and Titze, I. R. (2008). Vocal fold elasticity of the Rocky Mountain elk (*Cervus elaphus nelsoni*) -producing high fundamental frequency vocalization with a very long vocal fold. *J. Exp. Biol.* **211**, 2144-2154.

Riede, T., Lingle, S., Hunter, E. and Titze, I. R. (2010). Cervids with different vocal behavior demonstrate different viscoelastic properties of their vocal folds. *J. Morph.* **271**, 1-11.

Rowell, T. E. and Hinde, R. A. (1962). Vocal communication by the rhesus monkey *Macaca mulatta*, *Proc. R. Soc. Lond. B. Biol. Sci.* **138**, 279-294.

't Hart, J., Collier, R. and Cohen, A. (1990). *A Perceptual Study of Intonation.* Cambridge, MA: Cambridge University Press.

Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *J. Acoust. Soc. Am.* **83**, 1536-1552.

Titze, I. R. (2000). *Principles of Voice Production.* Salt Lake City, Utah: National Center for Voice and Speech.

Titze, I. R. (2006). *The Myoelastic-Aerodynamic Theory of Phonation.* Salt Lake City, Utah: National Center for Voice and Speech.

Titze, I. R. (2008). Nonlinear source-filter coupling in phonation: theory. *J. Acoust. Soc. Am.* **123**, 2733-2749.

Titze, I. R., Jiang, J. J. and Lin, E. (1997). The dynamics of length change in canine vocal folds. *J. Voice* **1**, 267-276.

Titze, I. R., Riede, T. and Popollo, P. (2008). Vocal exercises to determine nonlinear source-filter interaction. *J. Acoust. Soc. Am.* **123**, 1902-1915.

Tsai, C., Shau, Y. W., Liu, H. M. and Hsiao, T. Y. (2008). Laryngeal mechanisms during human 4-kHz vocalization studied with CT, videostroboscopy, and color Doppler imaging. *J. Voice* **22**, 275-282.

Ufema, K. and Montequin, D. (2001). The Performance Scream: vocal use or abuse? In *The Voice in Violence and other Contemporary issues in Professional Voice and Speech Training* (ed. R. Dal Vera), pp. 74-86. Cincinnati, OH: Voice and Speech Trainer Association, Incorporated.

van den Berg, J. W. (1955). On the role of laryngeal ventricle in voice production. *Folia Phoniatrica* **7**, 57-69.

van den Berg, J. W. (1958). Myoelastic-aerodynamic theory of voice production. *J. Speech Hear. Res.* **1**, 227-243.

West, R. and Larson, C. R. (1993). Laryngeal and respiratory activity during vocalization in macaque monkeys. *J. Voice* **7**, 54-68.

Zhang, K., Siegmund, T., Chan, R. W. and Fu, M. (2009). Predictions of fundamental frequency changes during phonation based on a biomechanical model of the vocal fold lamina propria. *J. Voice* **23**, 277-282.