Frequencies, bandwidths and magnitudes of vocal tract and surrounding tissue resonances, measured through the lips during phonation

Noel Hanna,[a] John Smith, and Joe Wolfe

School of Physics, University of New South Wales, Sydney, New South Wales 2052, Australia

The frequencies, magnitudes, and bandwidths of vocal tract resonances are all important in understanding and synthesizing speech. High precision acoustic impedance spectra of the vocal tracts of 10 subjects were measured from 10 Hz to 4.2 kHz by injecting a broadband acoustic signal through the lips. Between 300 Hz and 4 kHz the acoustic resonances R (impedance minima measured through the lips) and anti-resonances R̄ (impedance maxima) associated with the first three voice formants, have bandwidths of $\sim$50 to 90 Hz for men and $\sim$70 to 90 Hz for women. These acoustic resonances approximate those of a smooth, dry, rigid cylinder of similar dimensions, except that their bandwidths indicate higher losses in the vocal tract. The lossy, inertive load and airflow caused by opening the glottis further increase the bandwidths observed during phonation. The vocal tract walls are not rigid and measurements show an acousto-mechanical resonance R0 $\sim$ 20 Hz and anti-resonance R̄0 $\sim$ 200 Hz. These give an estimate of wall inertance consistent with an effective thickness of 1–2 cm and a wall stiffness of 2–4 kN m^{-1}. The non-rigidity of the tract imposes a lower limit of the frequency of the first acoustic resonance f_{R1} and the first formant $F1$.
© 2016 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4948754]

I. INTRODUCTION

Many of the acoustic features that distinguish speech phonemes are determined by resonances of the vocal tract. The resonances of the tract as viewed from the vocal folds act to match the high impedance of the glottis to the low impedance of the radiation field at the lips. Thus these resonances produce peaks in the spectral envelope outside the mouth, which are known as formants (Fant, 1970).[1] Formants and how they vary in time characterize vowels and are important to some consonants.

The bandwidth B of these resonances provides valuable information for speaker identification (Childers and Wu, 1991; Hanson and Chuang, 1999), and about the glottal source (Childers and Wu, 1991; Flanagan, 1972; Rothenberg, 1981) and may provide useful information for the diagnosis of obstructive sleep apnea (Robb et al., 1997). The bandwidth is often normalized and expressed as the quality factor $Q = f/B$, where f is the frequency. Wide B, and therefore low Q resonances, produce weaker formants, so that sounds may not be recognizable as distinct vowels, whereas narrow B resonances increase intelligibility (Summerfield et al., 1985). However, losses in the human tract will presumably place an upper limit on Q, and indeed synthesized voices with high Q resonances sound unnatural.

Previous data concerning the resonances obtained from the radiated voice sound give an estimation of B with a resolution comparable with the fundamental frequency f_o (e.g., Peterson and Barney, 1952). Higher frequency resolution has been achieved with swept sine excitation at the neck with the glottis held closed (Van den Berg, 1955; Dunn, 1961; Fant, 1972) and held open (Fujimura and Lindqvist, 1971). Measurements during phonation have been demonstrated using a broadband vibration applied outside the neck (Pham Thi Ngoc and Badin, 1994). However, these measurements are subject to transfer functions at the neck that are not measured simultaneously, so that the actual frequency response of the injected sound in the vocal tract is unknown. Neck related transfer functions have been measured independently, however (Meltzner et al., 2003; Wu et al., 2014). Attempts to provide an impulsive signal by popping a balloon inside the tract overcome this issue but have a low signal-to-noise ratio and have not yielded consistent results (Russell and Cotton, 1959; Coffin, 1989). A microphone inserted via the nasal cavity to record the sound close to the glottis is uncomfortable and the frequency response is sensitive to the position. A preliminary determination of the vocal tract transfer function in this way did not yield an improved spectral response (Kob, 2002). Measurements with external microphones and broadband excitation at the open mouth with an electric spark (Tarnóczy, 1962; House and Stevens, 1958) or a synthesized broadband signal (Epps et al., 1997) are made in parallel with the radiation impedance at the open mouth, so that the magnitude of the impedance is not accurately known and thus it is currently impractical to calculate bandwidths. Wodicka et al. (1990) input a broadband sound source through the lips and monitored the response of the supra and subglottal vocal tract using accelerometers; this method, however, required breathing through the nose and thus a connection to the nasal tract.

The methods of Pham Thi Ngoc and Badin (1994) and Epps et al. (1997) allow measurements during phonation,

---

[a] Electronic mail: n.hanna@unswalumni.com

and these demonstrate that there are differences in bandwidth between the closed glottis gesture and phonation (Pham Thi Ngoc and Badin, 1994), and between resonance frequencies in whisper, creak, and normal voice (Swerdlin et al., 2010). However, the most commonly used bandwidth data for modeling the voice is based only on closed glottis measurements (Childers and Wu, 1991; Hawks and Miller, 1995), which therefore may not be representative of the bandwidths produced during speech or singing.

In this study, a cylindrical impedance head is used to inject a broadband signal into the mouth, with the subject's lips sealed around the measurement head. The measured response to this signal yields the frequencies, magnitudes, and bandwidths of the resonances during phonation, while avoiding the disadvantages detailed above. The impedance minima as seen from the lips are here called resonances Ri, where i denotes an index that increases with increasing frequency; these occur at frequencies close to those of the formants Fi. A limitation of this method is that sealing the lips around the impedance head restricts the frequency range of the first acoustic resonance, f_{R1}, hence the speech sounds used are not phonetically meaningful. However, the magnitudes and bandwidths of the resonances from the measured impedance spectrum give good estimates of the losses, and the low frequency resonances give useful information about the wall mechanics.

Researchers have recognized that the mechanical properties of the tract would be important in modeling the tract at low frequencies (e.g., Wakita and Fant, 1978) and that any mechanical resonances are likely to be important for the production of plosives. Ishizaka et al. (1975) used mechanical vibration to excite the cheek and neck and found broad resonances in the range 30–70 Hz. Fant et al. (1976) excited the tract acoustically and found maximum tissue vibration around 200 Hz, similar to the frequency of the lowest formant found when phonation was studied under increased ambient air pressure conditions (Fant and Sonesson, 1964; Fujimura and Lindqvist, 1971). However, each of these studies considered only a single acousto-mechanical resonance in isolation.

The present study measures the acoustic and acousto-mechanical resonances and anti-resonances simultaneously from the response of the vocal tract to broadband acoustic excitation through the lips delivered by a three-microphone impedance head. It gives precise measurements of their frequency, amplitude, and bandwidth both without and during phonation. Using simple models, the magnitudes, frequencies, and bandwidths of the acoustic resonances are compared with those of a simple duct having comparable dimensions, and those of the mechanical resonances are related to the mechanical parameters of the relevant tissues.

II. MATERIALS AND METHODS

A. The subjects

Ten volunteer subjects (seven men and three women) took part in the study, whose design was approved by the university's human ethics committee. All subjects were non-smokers and none reported a history of voice disorders. The age and height of the subjects is given in Table I. Subject S2 is a professional bass baritone opera singer, S8 is an amateur soprano, all other subjects had received no formal training in speaking or singing. The experiments were conducted in a room treated to reduce reverberation and external noise.

B. Measuring acoustic resonances of the vocal tract

Acoustic impedance measurements were made following the three-microphone three-calibration method of Dickens et al. (2007). A cylindrical aluminum measurement head was constructed with internal diameter 26.2 mm and microphones spaced at 20, 60, and 260 mm from the measurement plane end. A broadband signal was synthesized from sine waves between 10 Hz and 4.2 kHz with a spacing of 2.69 Hz ($44.1 \, kHz/2^{14}$) and phases chosen to improve the signal-to-noise ratio (Smith, 1995). A single cycle of the signal thus lasts 370 ms. This compromise between time and frequency resolution allows a frequency resolution of ~ 3 Hz, while also allowing the stability of the phonation to be assessed by comparing successive cycles during a phonation that lasts several seconds. The acoustic signal was injected via an enclosed 2 in. dome midrange loudspeaker (Jaycar CM-2092, Australia) at the opposite end. The signals from the three microphones (Brüel & Kjær 4944A, Denmark) were connected via a conditioning preamplifier (Nexus 2690, Brüel & Kjær, Denmark) to a FireWire audio interface (MOTU 828, Cambridge, MA) and sampled at 44.1 kHz with 16 bit resolution. The upper frequency limit of the measurements is determined by the smallest spacing between any two of the three microphones, which here causes a singularity at 4.3 kHz. The lower frequency limit of around 10 Hz was determined by the frequency response of the loudspeaker, microphones, and interface.

The impedance head was calibrated using three non-resonant acoustic loads: a quasi-infinite impedance (a large rigid mass); a resistive impedance (an acoustically infinite pipe); and a large flange (Dickens et al., 2007). Measurements from the microphone array determine the acoustic impedance Z at a measurement plane at the end of the cylindrical impedance head located 10 mm inside each subject's lips, as shown in the schematic of the apparatus in Fig. 1.

TABLE I. Details of experimental subjects.

Men	Age (years)	Height (m)
S1	57	1.83
S2	55	1.84
S3	31	1.84
S4	33	1.71
S5	63	1.65
S6	60	1.78
S7	31	1.84
Mean ± S.D.	47 ± 15	1.78 ± 0.08
Women		
S8	22	1.70
S9	31	1.77
S10	31	1.74
Mean ± S.D.	28 ± 5	1.74 ± 0.04

J. Acoust. Soc. Am. **139** (5), May 2016

Hanna *et al.*   2925

With this measurement technique, the geometry on the upstream (loudspeaker) side of the microphones does not affect the measurements. Here, for the measurements on subjects, a narrow tube (5 cm long and 7.8 mm diameter) was connected in parallel with the loudspeaker and filled with acoustic fiber to permit the direct current (DC) airflow necessary for phonation. The impedance head was disinfected with ethanol after each experimental session.

With the exception of preliminary reports (Hanna *et al.*, 2013; Wolfe *et al.*, 2013), this technique has not been evaluated in the presence of DC airflow. So it is important to clarify the effects of airflow, turbulence, gas composition, and temperature on the measured impedance spectra. Both the frequency and magnitudes of extrema in acoustic impedance have been shown to depend weakly on breath temperature and the concentrations of water and CO_2 in the breath (Boutin *et al.*, 2013).

In order to quantify these effects, measurements were made on a simplified model system. A rigid plastic cylinder with the same inner diameter as the measurement device was closed at the far end with the termination penetrated by a "glottis" of 1 mm radius and 10 mm length at a length 175 mm from the measurement plane.

The impedance of the cylinder was measured in the range 200–4200 Hz at room temperature (25 °C, 58% humidity) immediately after calibrating the measurement device. To estimate the effects of temperature, the cylinder was then placed in a waterproof plastic bag and heated to 37 °C in a controlled temperature bath (NB1-80-T, Thermoline, Australia). The dry, warm cylinder was measured again, using the impedance head whose mass (1.4 kg) kept it within 1 °C of room temperature throughout this and later experiments.

From 25 °C to 37 °C (298 to 310 K), the measured resonance frequencies increase by 2%. Due to the temperature alone, the expected change due to the increased speed of sound is an increase of $\sqrt{310/298} - 1 = 2.0\%$. Over the same temperature increase, the extrema in magnitudes of impedance varied by as much as 10%, sometimes increased, sometimes decreased. This is due to the sharpness of the peaks in this model system, their shifts, and the finite sampling in the frequency domain. This represents the measurement precision of the system in this situation. An increase in relative humidity from 58% to a level whereby water droplets condensed on the walls produced only negligible changes in the resonance frequencies and magnitudes.
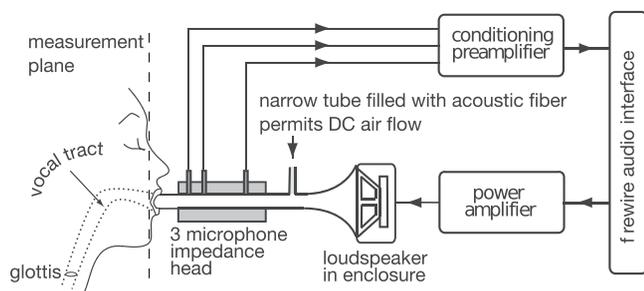

FIG. 1. Schematic diagram (not to scale) showing how the vocal tract impedance was measured either during phonation or with the glottis closed.

To estimate the effects of airflow alone, without changes in the gas composition, the cylinder (at room temperature) was measured while a supply of compressed air was blown through the glottal aperture at rates of 4 L s^{-1} to 10 L s^{-1}. For flows up to 6 L s^{-1}—a value near the upper limit of human phonation—the measured f_{Ri} change by less than 1%, which is the limit of frequency resolution for f_{R1} in this configuration. The net effect of frequency shifts due to steady airflow without turbulence is expected to be small. For example, 5 L s^{-1} gives a flow velocity v of order 5×10^{-3} m^3 s$^{-1}/\pi$ (radius^2) m$^2 = 9$ m s^{-1} and thus an expected frequency shift (due to non-cancellation of the Doppler shifts in the two directions) of $1 - 1/\sqrt{1 - v^2/c^2} = 0.04\%$. In the presence of airflow of 10 L s^{-1}, the measured f_{Ri} decreases by 2% and the noise on the measured impedance spectrum increases due to turbulence. This noise partially obscures the impedance extrema and may cause the impedance magnitude to be slightly underestimated and hence the bandwidths to be overestimated during very loud phonation.

Finally, to estimate the combined effects of temperature, gas composition, and airflow, an experimenter blew air through the glottal aperture toward the measurement head. These measurements were repeated approximately every minute for 20 min, then the cylinder was allowed to return to room temperature. With all effects in combination, the measured f_{R1} decreases by 1% and the changes in f_{Ri} for $i > 1$ are all less than 1%. Boutin *et al.* (2013) also found that the net change in the speed of sound in exhaled air is small, with the fall due to increased CO_2 tending to compensate for the rises due to temperature and humidity.

These results indicate that the measurements using 37 °C humid air and in the presence of airflow (analogous to the measurements made on live subjects) with a room-temperature impedance head give values accurate to ±1% in frequency and ±10% in magnitude. In the results described later, this uncertainty in magnitude produces an uncertainty of typically 10% in calculated bandwidth. These uncertainties are considerably less than the variation expected between subjects, and even for repeated measurements on a single subject.

C. Experimental protocol

The subjects were asked to find a comfortable position for the impedance head in the mouth and to ensure an airtight seal with their lips around the outer diameter of 31.8 mm. For phonation the subjects were asked to produce a sustained vowel with the velum closed and with the tongue in the position to pronounce the word "heard," for the duration of the broadband signal. This vowel sound in Australian English corresponds approximately to the neutral vowel [ɜː] in the International Phonetic Alphabet. However, the tract configuration will only roughly approximate that of that vowel when spoken, so no phonetic significance of the resonance frequencies is implied. For measurements with the glottis closed, subjects were asked to phonate and then abruptly stop while retaining the same vocal gesture. Subjects were given time to practice these unusual configurations, and some subjects chose to pinch their nose to check

that they were not nasalizing. At least three measurements were made for each subject during (a) a comfortable, low-pitched, modal phonation comparable with their speaking voice, and (b) while miming the same gesture with the glottis closed.

During each gesture, 10 or 18 cycles of the broadband signal were injected through the measurement head into the subject's mouth. The first and last cycles of the measured signal were discarded to exclude acoustic transients and to exclude the possible influence of the carrier word. The measured impedance for each cycle was displayed to the experimenter immediately following the recorded cycles so that changes in the geometry of the vocal tract during measurements could be detected and the measurement repeated if necessary. Later, the approximate location of the maxima and minima of the impedance spectra were identified with a sixth order Savitzky–Golay smoothing filter (Savitzky and Golay, 1964). Parabolas were fitted to both sides of each of the extrema and their crossover point used to determine the frequency f and the impedance magnitude $|Z|$. The bandwidth B was determined as the frequency range at half maximum power.

D. Vocal tract properties at low frequency and during phonation

Measurements with the glottis closed were initially made using a single broadband spectrum from 14 to 4200 Hz. However, the low frequency components in the initial broadband signal excited acousto-mechanical resonances in the tissues surrounding the vocal tract, and this excitation sometimes disturbed some of the subjects and caused them to alter their tract geometry during a measurement, or to produce an uncontrolled vibrato. In either case, this interfered with the measured impedance spectrum. For this reason, the frequency range of interest was also divided into three different broadband frequency ranges. To determine the acoustic resonances, both with glottis closed and during phonation, a broadband signal of 200–4200 or 300–4200 Hz with a resolution of 2.69 Hz ($44.1\,\text{kHz}/2^{14}$) was injected.

To study the acousto-mechanical resonances, two narrower, low frequency windows were used: 10–50 Hz with a frequency resolution of 0.34 Hz ($44.1\,\text{kHz}/2^{17}$), and 14–300 or 14–400 Hz with a resolution of 0.67 Hz ($44.1\,\text{kHz}/2^{16}$). During these measurements, four of the subjects also had a small magnet attached to the cheek and/or neck to map the surface vibration by inducing a voltage in a coil of wire placed on the normal axis a fixed distance away.

E. Simple model of vocal tract as rigid cylinder

Calculations using a simple cylindrical model of the tract can provide a useful point of comparison with acoustic impedance measurements. For frequencies below about 4 kHz, wavelengths λ are much greater than the transverse dimensions of the vocal tract, so a simple one-dimensional approximation of the tract as a rigid tube captures much of the physics. Curvature of such a tube has little effect in the frequency range studied (Sondhi, 1986) and is neglected in this analogy. At high frequencies, deviations from cylindrical geometry can become more important and, at

frequencies below about 200 Hz, the non-rigidity of the walls must be considered (Sondhi, 1974; Makarov and Sorokin, 2004) as discussed in Sec. III A 4.

The impedance Z of a rigid cylinder loaded by an impedance Z_L is given by the following relation (Fletcher and Rossing, 1998):

$$Z = \frac{\rho c}{\pi r^2} \left(\frac{Z_L \cos(\Gamma l) + j \frac{\rho c}{\pi r^2} \sin(\Gamma l)}{j Z_L \sin(\Gamma l) + \frac{\rho c}{\pi r^2} \cos(\Gamma l)} \right), \tag{1}$$

where ρ is the density of air, c is the speed of sound ($354\,\text{m s}^{-1}$ in saturated air at $37\,°\text{C}$), r is the radius, l is the length, and Γ is the complex wave number

$$\Gamma = \frac{\omega}{c} - j\alpha \approx \frac{\omega}{c} - j \frac{1.2 \times 10^{-5} \sqrt{\omega}}{r}, \tag{2}$$

where ω is the angular frequency and α is an attenuation coefficient (Fletcher and Rossing, 1998), which accounts for losses to the tube walls by viscous drag and thermal conduction.

The approximately equal spacing of the formants of the neutral vowel [ɜː] is approximately consistent with those of a cylinder of appropriate length closed at the glottis end, i.e., $Z_L = \infty$, which gives

$$Z = -j \frac{\rho c}{\pi r^2} \cot(\Gamma l). \tag{3}$$

Setting the glottal closure as complete and setting $\alpha = 0$ for the following estimation, the impedance minima, i.e., the resonances Ri that fall near the formants Fi, occur when $\lambda_{Ri} = 4l$, $4l/3$, $4l/5$, etc. Impedance maxima, $\bar{R}i$, occur when $\lambda_{\bar{R}i} = 2l$, $2l/2$, $2l/3$, etc. An effective vocal tract length can then be computed as the arithmetic mean of the length determined by these relationships as follows:

$$l = \frac{1}{7} \left(\frac{c}{4f_{R1}} + \frac{c}{2f_{\bar{R}1}} + \frac{3c}{4f_{R2}} + \frac{c}{f_{\bar{R}2}} + \frac{5c}{4f_{R3}} + \frac{3c}{2f_{\bar{R}3}} + \frac{7c}{4f_{R4}} \right). \tag{4}$$

It is then possible to adjust the radius r of the cylinder and the attenuation coefficient α to be comparable with measured acoustic impedances.

III. RESULTS AND DISCUSSION

A. Resonances of the vocal tract with closed glottis

1. Impedance spectrum

Figures 2 and 3 show typical acoustic impedance spectra for men and women measured through the lips with the glottis closed. The minima above about 400 Hz are the acoustic resonances Ri that fall close to the formants Fi in speech. The maxima are anti-resonances $\bar{R}i$. For each resonance in those figures, the centre of the ellipse is indicated by a cross and shows the average data (x,y) for that resonance (and that sex). The slope of the long axis of each ellipse indicates the correlation between the (x,y) data for that resonance. The
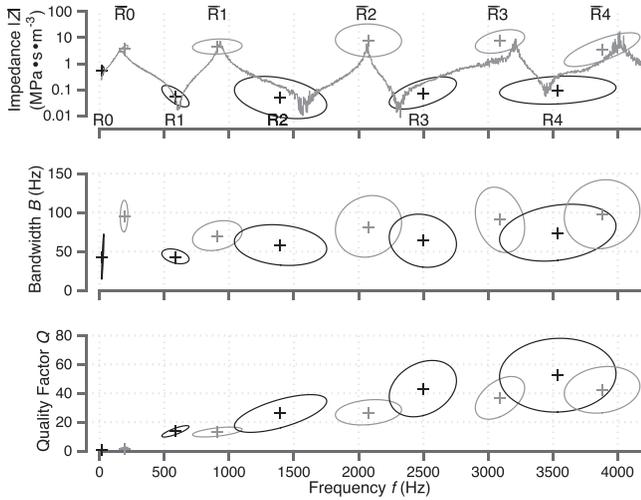
FIG. 2. The measured dependence on frequency of the impedance magnitude $|Z|$, bandwidth B, and Q factor for each resonance Ri (dark) and anti-resonance $\bar{\text{R}}$i (pale) for men, with the glottis closed and measured through the lips. An example impedance spectrum is included in the impedance plot. The ellipse semi-axes indicate one standard deviation.

lengths of the semi-axes represent one standard deviation calculated in that direction for all subjects of that sex.

In speech, the frequency of the first formant $F1$ typically has a value between about 300 and 800 Hz that increases with mouth opening. The frequency of the second formant $F2$ varies between about 800 and 2000 Hz, and depends largely on tongue shape (Peterson and Barney, 1952). When the neutral vowel [ɜː] is pronounced by a man with vocal tract length 170 mm, then $f_{R1} \sim F1 \sim 500$ Hz, $f_{R2} \sim F2 \sim 1500$ Hz, etc. The mean f, B, and $|Z|$ of the resonances measured here are given in Table II. Note that in these experiments, the impedance head constrains the mouth opening for all subjects, limiting the possible range of f_{R1}. Also, the measurement plane was located about 10 mm inside the lips and thus the effective vocal tract length is shortened (see Fig. 1). Correction for this, and the absence of the external
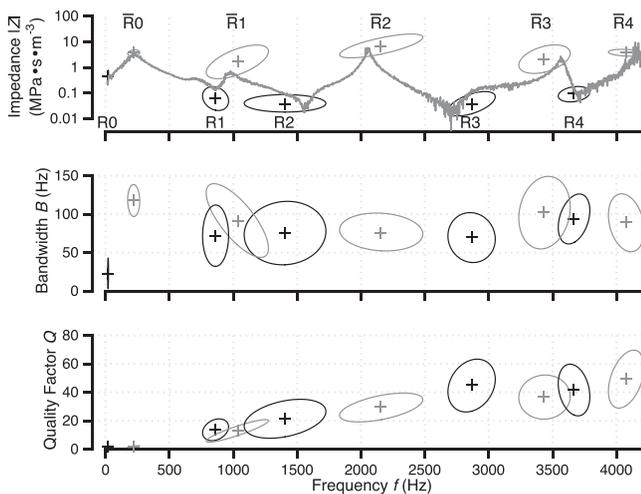

FIG. 3. The measured dependence on frequency of the impedance magnitude $|Z|$, bandwidth B, and Q factor for each Ri (dark) and $\bar{\text{R}}$i (pale) for women, with the glottis closed. An example impedance spectrum is included in the impedance plot.

radiation field at the lips, suggests that comparison with $F1$ requires f_{R1} to be decreased by a factor of approximately 15%–20% for men. However the comparison is more complicated and is discussed further in Sec. III B 3.

Although subjects were asked to keep the tongue low, the higher resonance frequencies typically displayed greater variation than those of the lower resonances. Presumably, small unintentional changes in the vocal tract geometry, as might happen with a slight difference in tongue position between measurements, would have a greater effect at high frequencies. (Any large tongue movement that occurred within a single measurement would be apparent and those data could be discarded—see Sec. II C.)

2. Comparing the vocal tract with a rigid cylinder

A simple cylindrical model of the tract, with an effective vocal tract length l calculated using Eq. (4) with the measured f_{Ri} and $f_{\bar{R}i}$ in Table II, is shown in Fig. 4. This provides a useful comparison with the measurements presented in Figs. 2 and 3 and Table II.

Equation (4) yields average effective lengths of 173 and 155 mm, respectively, for the men and women in this study, not including the ~10 mm of the vocal tract between the lips that lies beyond the measurement plane.

This effective length is only an approximation to the anatomical length due to the simplicity of the model. It is difficult to find accurate information on comparable geometrical vocal tract lengths but the calculated differences between sexes are comparable with the reported 10% difference (Fitch and Giedd, 1999), and the average length for men is about 10 mm longer than the 175 mm reported by Story et al. (1996) for a similar vowel in the word "earth," for a 1.70 m tall man (shorter than the average of 1.78 m in this study). Some difference is expected between effective and geometrical length, due to the approximation of the simple cylindrical model, as suggested by Fig. 4.

Three of the men and two of the women were comfortably able to put their teeth and lips around the impedance head, the remaining subjects sealed it with their lips only. No systematic difference was observed in the overall form of the impedance spectra and the subjects kept their chosen position throughout the experiment, with the standard deviation of each individual subject's f_{R1} typically less than 10%.

The calculated acoustic impedance spectrum shown in Fig. 4 is for a rigid-walled cylinder that is closed at the far (glottis) end, with length $l = 173$ mm and radius 10 mm (values close to the average male data) calculated using Eq. (3), assuming dry, rigid walls and the standard value of $\alpha = 1.2 \times 10^{-5}\sqrt{\omega}/r$ [Eq. (2)]. A qualitative difference between this simple model and the vocal tract is that, for frequencies below the first acoustic resonance f_{R1}, the impedance of a closed rigid tube increases with decreasing frequency—monotonically to infinity. However, the vocal tract is not rigid, so the resonance and anti-resonance that occur below f_{R1} in the measured impedance spectra can be attributed to mechanical properties of its walls, as discussed in Sec. III A 4.

TABLE II. The effect of phonation on the measured resonances. The mean resonance (R) and antiresonance (R̄) data for men (upper) and women (lower) with a closed glottis are shown in normal type. Measurements during phonation italicized. Phonation values marked with an asterisk are significantly different from values with glottis closed at the $p < 0.01$ level. Values for R̄4 are not included since it was not always within the measured frequency range. The number of samples is the number of cycles that displayed a clear minimum or maximum in the impedance spectrum associated with that resonance or anti-resonance. Not all extrema could be clearly identified in each cycle of each measurement. Data shown as mean ± standard deviation.

Men	R0	R̄0	R1	R̄1	R2	R̄2	R3	R̄3	R4
f (Hz)	25 ± 5	210 ± 55	585 ± 70	940 ± 155	1415 ± 250	2075 ± 170	2465 ± 190	3105 ± 285	3535 ± 295
	25 ± 10	*245 ± 35**	*570 ± 55*	*1010 ± 175**	*1390 ± 200*	*2015 ± 175**	*2425 ± 120*	*3005 ± 215**	*3620 ± 370*
B (Hz)	40 ± 20	105 ± 30	45 ± 5	70 ± 10	60 ± 15	80 ± 25	65 ± 20	100 ± 35	75 ± 25
	20 ± 10*	*85 ± 30**	*85 ± 30**	*125 ± 30**	*90 ± 25**	*100 ± 35**	*85 ± 20**	*105 ± 35*	*95 ± 20**
$\|Z\|$ (MPa s m^{-3})	0.55 ± 0.15	3.9 ± 1.0	0.069 ± 0.051	8.3 ± 10	0.12 ± 0.15	12 ± 11	0.13 ± 0.12	11 ± 0.65	0.13 ± 0.12
	0.39 ± 0.12*	*2.6 ± 0.96**	*0.15 ± 0.11**	*4.2 ± 4.6**	*0.17 ± 0.17**	*6.2 ± 3.5**	*0.22 ± 0.23**	*11 ± 7.9*	*0.18 ± 0.13**
No. samples	184	250	221	229	226	218	221	214	192
	100	*188*	*113*	*146*	*119*	*142*	*138*	*138*	*90*
Women	R0	R̄0	R1	R̄1	R2	R̄2	R3	R̄3	R4
f (Hz)	20 ± 5	215 ± 30	860 ± 65	1035 ± 160	1405 ± 210	2160 ± 215	2890 ± 150	3445 ± 135	3710 ± 150
	20 ± 5	*245 ± 25**	*800 ± 65**	*970 ± 165**	*1515 ± 95**	*2340 ± 125**	*2625 ± 180**	*3495 ± 145*	*3955 ± 200**
B (Hz)	30 ± 15	115 ± 15	70 ± 25	90 ± 30	75 ± 25	75 ± 15	75 ± 25	100 ± 30	95 ± 20
	15 ± 5*	*105 ± 35**	*75 ± 30*	*110 ± 45**	*90 ± 25**	*115 ± 35**	*90 ± 20**	*100 ± 25*	*90 ± 20*
$\|Z\|$ (MPa s m^{-3})	0.48 ± 0.18	3.7 ± 0.7	0.077 ± 0.057	2.4 ± 1.7	0.045 ± 0.024	9.0 ± 7.3	0.077 ± 0.12	0.29 ± 0.25	3.8 ± 0.71
	0.43 ± 0.11	*2.7 ± 0.8**	*0.28 ± 0.16**	*1.5 ± 0.6**	*0.071 ± 0.063**	*4.0 ± 3.4**	*0.18 ± 0.06**	*9.3 ± 5.8**	*0.31 ± 0.22**
No. samples	47	121	90	97	118	121	97	107	76
	22	*52*	*43*	*73*	*71*	*77*	*80*	*71*	*51*

3. Bandwidth and quality factor

Table II and Figs. 2 and 3 display the mean bandwidth B calculated for the experimental subjects. (Bandwidth measurements include an error of ±10% due the resolution of the magnitudes of impedance peaks: Sec. II B.) For measurements with the glottis closed, the bandwidths of the acoustic resonances and anti-resonances (those above 300 Hz) increase with increasing resonance frequency. Over the frequency range measured, the rate of this increase is roughly $\Delta B_{\mathrm{Ri}} / \Delta f_{\mathrm{Ri}} = 1$ Hz/100 Hz (bandwidth varying from about 50 to 90 Hz) for men and roughly 0.5 Hz/100 Hz ($B \sim 70$ to 90 Hz) for women.

For both impedance maxima and minima, the calculated bandwidths of a rigid tube are significantly narrower and the Q factors higher than those of the closed glottis vocal tracts (note the sharper extrema of the model curve in Fig. 4). For a rigid cylinder, the losses that determine B are dominated by visco-thermal energy losses at the boundary layer, represented by the attenuation coefficient α, which modifies the wavenumber, as shown in Eq. (2). The attenuation coefficient in the vocal tract is expected to be higher than that in a rigid, dry cylinder of comparable dimensions for reasons that include (i) it has larger surface area:volume ratio, (ii) it is not completely rigid, and (iii) it has a wet surface. [A wet surface on a rigid pipe, 19 mm in diameter, has been shown to increase losses and thus to increase bandwidths by a few percent (Coltman, 2003).] These loss mechanisms have two related effects: they increase the bandwidths B for both resonances Ri and anti-resonances R̄i, and they also decrease the magnitude $|Z|$ of maxima in impedance (anti-resonances when measuring through the lips) and increase the $|Z|$ of minima.

The effect of these energy losses can be simulated empirically by adjusting α to improve the agreement between the calculated model and the measured vocal tract. Figure 5 shows the calculated impedance and associated bandwidths of the model when α is increased by a factor of 5; the extrema then have magnitudes and bandwidths comparable with the measured data for men in Table II. The model used in Fig. 5 also includes non-rigid walls. This has only a minor effect on the impedance curve in the range of the acoustic resonances.
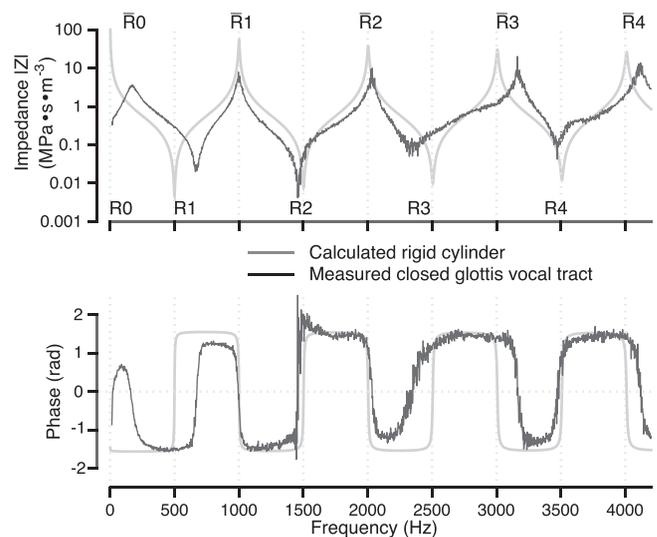

FIG. 4. Typical impedance spectrum (magnitude and phase) measured through the lips for 1 male subject over the range 14–4200 Hz, with the glottis closed (dark), and the calculated impedance of a closed rigid cylinder of comparable dimensions (pale). The measurements took 370 ms. [The low frequency data (10 to 400 Hz) are shown on an expanded frequency scale in Fig. 7.]

J. Acoust. Soc. Am. **139** (5), May 2016
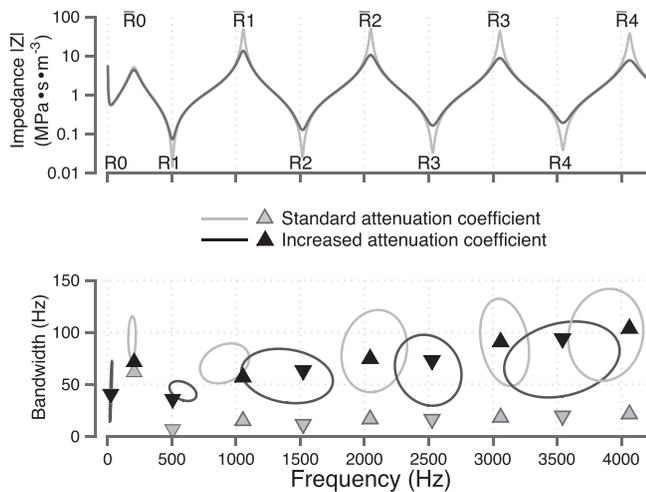
Hanna *et al.* 2929

FIG. 5. Calculated impedance magnitudes and bandwidths of the non-rigid, cylindrical vocal tract model from 1 to 4200 Hz, with the standard value of the visco-thermal attenuation coefficient α (pale line and pale triangles), and with α increased by a factor of 5 (dark line and dark triangles). The ellipses show the means and standard deviations of measured bandwidths for men from Fig. 2.

4. Acousto-mechanical resonances and the low frequency model

At frequencies below f_{R1}, two additional low frequency extrema, R0 and R̄0, are present in the impedance spectra measured through the lips, see Figs. 2 and 3, 4 and 5. These do not appear in the spectrum of a rigid tube (Fig. 4), but can be included in the simple model by including the finite mass and rigidity of the walls, with a few assumptions (Fig. 5). Again, the vocal tract is very simply modeled by a uniform cylinder with walls of surface area S, mass m, uniform thickness w, and a uniformly distributed spring constant k_t discussed below.

At frequencies well below f_{R1}, the dimensions of the tract are much less than the wavelength $λ$, so that the tract may be approximated as an acoustically compact object, i.e., one in which the pressure is approximately uniform. So, for these frequencies, pressure acting on the surface area S of the tract walls accelerates tissue with an inertance $L_t = m/S^2$ in series with an acoustic compliance $C_t = S^2/k_t$ due to the deformability (here assumed elastic) of the walls. The two elements are in series because the pressure-induced vibration of the walls moves both. The spring constant $k_t = 2\pi rl\, dP/dr$ for radial dilation dr due to a homogeneous pressure change dP. Additionally, since an increase in internal pressure also compresses the air within the tract, the compliance C of that air acts in parallel with the series circuit just mentioned, and can be represented mechanically as a spring with constant $k = S^2 C$. The two compliances, C and C_t, are in parallel at low frequency because an injected acoustic volume caused by a pressure dP in the tract equals $CdP + C_t dP$, where the first term represents the compression of the air in the tract and the second the dilation of the tract.

Figure 6 shows the measured impedance spectrum for the same male subject as Fig. 2 in the low frequency range. For comparison, a measurement made by the same subject during phonation is included. The measurements made
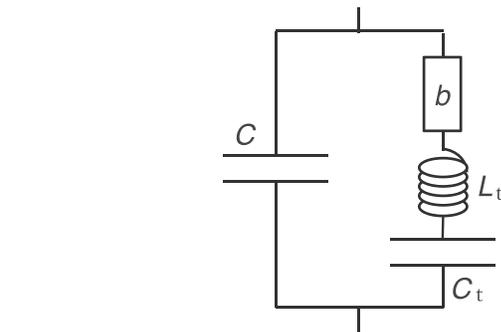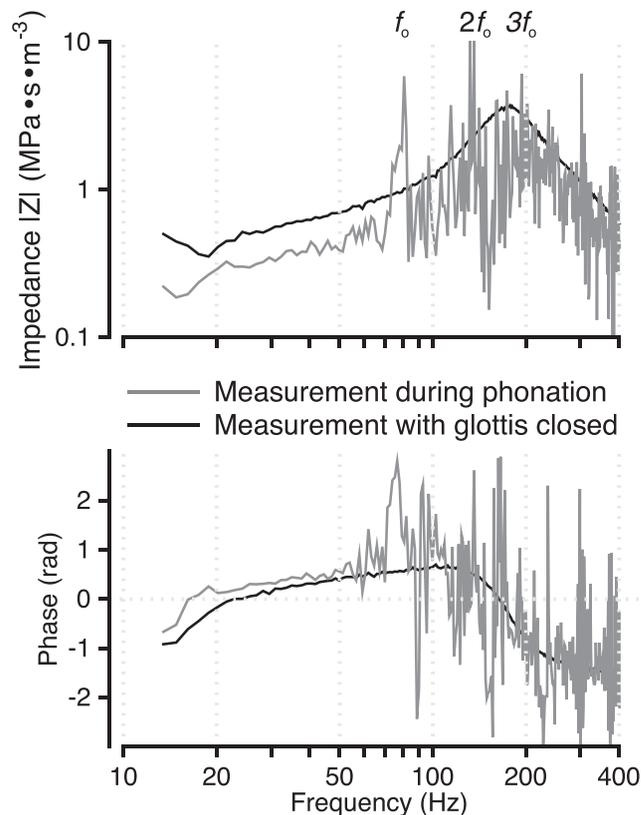

FIG. 6. (Top) Impedance spectra plotted on a log frequency scale showing the low frequency acousto-mechanical resonances for closed glottis condition (dark) and phonation (pale) for the same male subject. f_o, $2f_o$, and $3f_o$ label harmonics. (Bottom) Equivalent electrical circuit.

during phonation display more noise due to the DC airflow present. The low frequency acousto-mechanical resonance R0 indicated by the impedance minimum and zero phase crossing at about 20 Hz was close to the low frequency limit of the measurements (about 10 Hz), but was evident in the measurements on all subjects for both the closed glottis and phonation conditions.

As a simple model for the losses accompanying the tissue motion, a resistor b is positioned in series with L_t. Use of a resistor retains the linearity of the model. However, it is unlikely that the losses in the tissue are accurately modeled by a purely viscous term.

Including these mechanical elements and parameters into the simple vocal tract model introduces the additional low frequency series acousto-mechanical resonance R0, and parallel anti-resonance R̄0, in addition to the purely acoustic resonances R1,2,3,... and anti-resonances R̄1, 2, 3,...

Electrical analogues have previously been proposed to describe the expected acousto-mechanical resonances of the tract (a summary of this approach is given by Wakita and Fant, 1978). However, they have typically lacked some of the parameters necessary to define the system completely. Makarov and Sorokin (2004) compared calculations to show that using uniform values for the wall parameters is no less accurate for the purpose of determining formant frequencies than separating the tract into regions of varying wall mass, thickness, etc.

Since the measurements show $f_{\bar{R}0} \sim 10 f_{R0}$, the anti-resonance $\bar{R}0$ and resonance $R0$ can be treated separately for the purpose of estimating parameters: at low frequency (i.e., evaluating $R0$), C is approximately an open circuit, while at high frequency (i.e., evaluating $\bar{R}0$) C_t is effectively a short circuit. Because the same tissue mass is presumed to move in the two cases, the same wall inertance L_t is used for both the parallel and series resonances. The mechanical properties can then be estimated from the measured values of f, Q, and $|Z|$.

For the acoustic impedance minimum due to the resonance produced by the series combination of C_t and L_t, $f_{R0} = 1/2\pi\sqrt{L_t C_t}$. For the resonance involving C and L_t, which are acoustically in parallel, $f_{\bar{R}0}$ is measured at the impedance maximum and is slightly lower than the characteristic frequency f of the circuit due to the finite Q. In practice, there is little difference between these frequencies; nonetheless the two frequencies are related in the following way (Langford-Smith, 1963):

$$f = \frac{1}{2\pi\sqrt{L_t C}} = f_{\bar{R}0}\Bigg/\sqrt{\frac{-1}{Q_{\bar{R}0}^2} + \sqrt{\frac{2}{Q_{\bar{R}0}^2} + 1}}. \qquad (5)$$

This characteristic frequency and the measured $|Z|_{\bar{R}0}$ determine C, since $|Z|_{\bar{R}0} = Q_{\bar{R}0}/2\pi f C$ (Langford-Smith, 1963), and hence L_t from Eq. (5).

These acoustic components are related to tissue parameters by geometry. The volume of air in the tract

$$V = C\gamma P_A, \qquad (6)$$

where γ is the adiabatic constant and P_A the atmospheric pressure. For an approximately cylindrical vocal tract, the effective radius $r = \sqrt{V/\pi l}$, where l is the effective length of the tract from Eq. (4). So, the surface area of the tract walls $S = 2\pi r l$, which can be used to determine the effective mass of the walls $m = S^2 L_t$, and their thickness $w = SL_t/\rho$, assuming a tissue density $\rho \sim 10^3 \text{ kg m}^{-3}$.

Returning to the series resonance $R0$, the compliance of the tissue C_t is readily calculated since $f_{R0} = 1/2\pi\sqrt{L_t C_t}$, which determines the tissue spring constant $k_t = S^2/C_t$. Finally, the resistance terms in both circuits are given by the relations

$$b_{\bar{R}0} = 2\pi f L_t/Q_{\bar{R}0} = \sqrt{L_t/C}/Q_{\bar{R}0}, \qquad (7)$$

$$b_{R0} = 2\pi f_{R0} L_t/Q_{R0} = \sqrt{L_t/C_t}/Q_{R0}. \qquad (8)$$

The values determined from these calculations are given in the first two columns of Table III.

The maximum in impedance $|Z|_{\bar{R}0}$ corresponds to the parallel resonance of L_t and C, i.e., the oscillation of the walls primarily on the "spring" of the compliance of the air in the tract. $Q_{\bar{R}0}$ is low, which is likely associated with the damping caused by tissue deformation rather than loss in the

TABLE III. (Left) The acousto-mechanical resonance parameters measured (upper part of the table in bold case) and derived (lower part of table) in this study. (Middle and right) At right, for comparison, are the values reported by Fant *et al.* (1976) and Ishizaka *et al.* (1975). Parameter values in parentheses indicate values that were not presented in the original reference and have been calculated using their original data and the assumptions discussed in this paper. The numbers of male and female subjects are given in parentheses.

| | This study | | Fant *et al.* | | Ishizaka *et al.* | Male (1) | | |
	Male (7)	Female (3)	Male (5)	Female (7)	Relaxed cheek	Tense cheek	Neck
f_{R0} (Hz)	**25**	**19**			32	60	72
$f_{\bar{R}0}$ (Hz)	**212**	**216**	191	218			
$|Z|_{R0}$ (MPa s m^{-3})	**0.55**	**0.48**			(0.72)	(0.95)	(2.1)
$|Z|_{\bar{R}0}$ (MPa s m^{-3})	**3.9**	**3.7**	3.0	4.5			
B_{R0} (Hz)	**41**	**28**			62	115	153
$B_{\bar{R}0}$ (Hz)	**104**	**116**	76	94			
Q_{R0}	**0.8**	**1.0**			0.52	0.52	0.47
$Q_{\bar{R}0}$	**2.1**	**1.9**	2.5	2.3			
L_t (kg m^{-4})	1.4×10^3	1.4×10^3	1.1×10^3	1.4×10^3			
C (m^3 Pa^{-1})	4.1×10^{-10}	3.7×10^{-10}	(6.9×10^{-10})	(3.7×10^{-10})			
V (m^3)	5.8×10^{-5}	5.2×10^{-5}	9.8×10^{-5}	5.3×10^{-5}			
l (m)	0.17	0.16					
r (m)	0.010	0.010					
S (m^2)	0.011	0.010					
m (kg)	0.17	0.15			0.23	0.17	0.27
w (m)	0.02	0.01					
C_t (m^3 Pa^{-1})	2.9×10^{-8}	4.9×10^{-8}					
k_t (kN m^{-1})	4.3	2.1			9.5	3.7	5.5
b_{R0} (MPa s m^{-3})	0.27	0.18			(0.72)	(0.95)	(2.1)
$b_{\bar{R}0}$ (MPa s m^{-3})	0.86	1.0					

J. Acoust. Soc. Am. **139** (5), May 2016

Hanna *et al.*   2931

spring of the air. The minimum in impedance $|Z|_{\bar{R}0}$ is the result of the series resonance of L_t and C_t, i.e., the tissue mass oscillating on its own stiffness. $Q_{\bar{R}0} \sim 1$, suggesting that the spring of the tissue vibration is strongly damped. The equivalent circuit shows only one lossy component b. If the resistance acted purely as a viscous loss then the loss terms in Eqs. (7) and (8) would be equal, i.e., $b_{R0} = b_{\bar{R}0}$. However, these values differ by a factor of ~ 3 for men and ~ 5 for women, which implies that the resistance term increases with frequency. This frequency dependence is not explained by considering the losses as viscous.

Figure 5 shows a model impedance spectrum calculated using the cylindrical vocal tract with increased attenuation coefficient from Sec. III A 3, with the men's L_t, C_t, and mean b from Table III included in parallel. This model now includes the acousto-mechanical resonances, with values of f, $|Z|$, and B chosen to be comparable with the measured impedance. Note that using either b_{R0} or $b_{\bar{R}0}$ rather than the mean of the two values would give better agreement with the associated $|Z|$ and B.

Vibrations at the acousto-mechanical resonances were perceptible by the subject when excited by the low frequency injected signal. Measurement of the voltage induced in a coil of wire by the magnet mounted on the skin showed that the distribution of vibration around the face and neck was broadly in agreement with the contour plot produced by Fant *et al.* (1976) using an injected acoustic excitation with a closed glottis. The mean frequency for maximum wall vibration in the Fant *et al.* study was reported to occur close to the $f_{\bar{R}0}$ found in the current study. These data are shown in Table III along with the estimated L_t and V. With these values, C can be inferred from Eq. (6). The results presented in this paper are also consistent with an earlier study on a single male subject (Fujimura and Lindqvist, 1971), and estimates of the lowest formant frequency under the elevated air pressure conditions experienced by divers (Fant and Sonesson, 1964; Fujimura and Lindqvist, 1971).

For R0, direct comparisons with the literature are less straightforward. The mechanical impedance of the cheek and neck of a single male subject was determined directly by external excitation with a shaker (Ishizaka *et al.*, 1975) and used to calculate the equivalent m, k, and b per unit area. These data are also compared with the present study in Table III under the assumption that the vibration affects the same vocal tract surface area S used here, and $Z_{\text{acoustic}} = Z_{\text{mechanical}}/S^2$. In all three excitation cases (relaxed cheek, tense cheek, and neck) f_{R0} is higher and B_{R0} is larger than those measured in the current study, perhaps due to differences in the different distributions of force and displacement when excitation by pressure in the tract is compared with external excitation using a shaker. Despite this considerable difference, the effective $|Z|$, m, k, and b values for the cheek are of the same order of magnitude as those presented here.

Finally the values of m, k, and b determined in this study are also comparable with the 0.12–0.23 kg, 0.013–13 kN m^{-1}, and 0.48–0.98 MPa s m^{-3}, respectively, quoted "according to various sources" by Makarov and Sorokin (2004).

B. Resonances of the vocal tract during phonation

1. Differences between closed glottis and phonation measurements

The rapid opening and closing of the glottis during phonation affects the measured impedance in several ways. Most conspicuously, the measured impedance spectra include superimposed signals at the fundamental frequency of oscillation f_o and its harmonics $2f_o$, $3f_o$, ..., with phase uncorrelated with the injected broadband signal, as shown in Fig. 7. Consequently the superposed signal can either add or subtract according to the relative phase. Figure 7 shows three impedance spectra measured in a contiguous sequence of eight measurement cycles (each lasting 370 ms) for one subject. In the first (top) he is phonating with $f_o = 88$ Hz. Superimposed over the acoustic impedance (the broadband signal) are the harmonics of the voice, which are visible up to ~ 2 kHz. $8f_o$ coincides with f_{R1} and this coincidence obscures the bottom of the impedance minimum. The middle plot shows the measurement cycle during which the phonation ceased, and the bottom plot shows the first cycle after phonation had ceased. The resonances R1–4 and anti-resonances $\bar{R}0 - 3$ are indicated on the middle plot. Note that the sharpness of the peaks increases (i.e., B decreases) when phonation ceases as a consequence of the closed glottis, and that the frequencies of some of the resonances change slightly, e.g., f_{R4} increases from phonation to closed glottis.

Disregarding the harmonics of the voice, proportional changes between data from the closed glottis gesture and low-pitched $M1$ phonation are summarized in Fig. 8. Their significance was assessed using an unpaired t-test and differences with $p < 0.01$ are indicated with filled bars in the figures (and with an asterisk in Table II). For the grouped male data, changes in the frequencies of the acoustic R_i are insignificant at the 1% level, i.e., the closed glottis configuration preserves the resonances (therefore vowel) of the phonation measurement. The frequency changes of the anti-resonances
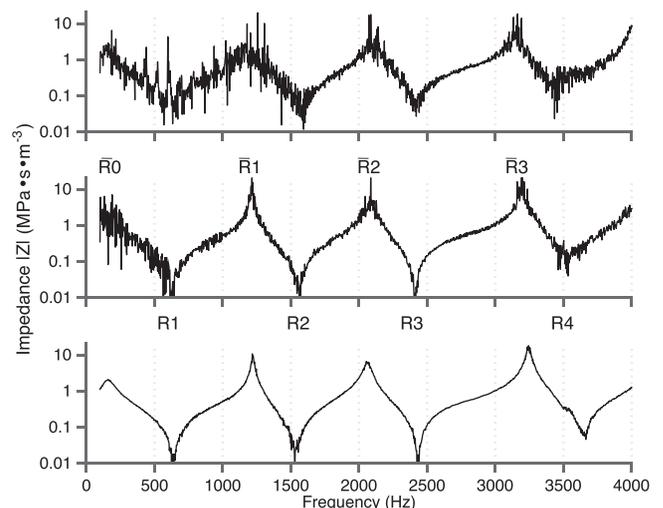

FIG. 7. Three measured impedance spectra (each measured over 370 ms) for one male subject, initially phonating (top). He stops phonating during the cycle shown in the middle. The bottom plot shows the first cycle after phonation had ceased.
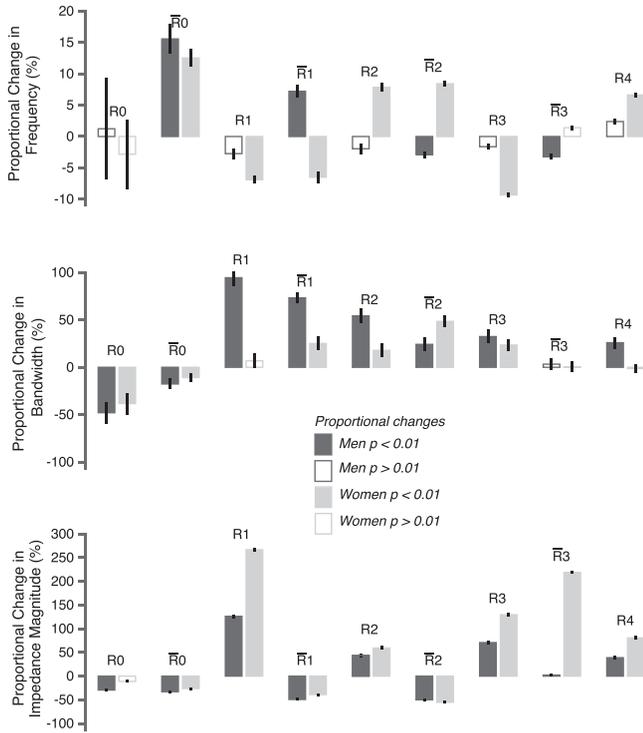
FIG. 8. Proportional changes in resonance frequency (top), bandwidth (middle), and impedance magnitude (bottom) when subjects change from closed glottis to M1 phonation for all subjects, i.e., [f(phonation)–f(closed)]/f(closed). Note the different scales. Filled bars indicate a significant change at the $p < 0.01$ level for men (dark) and women (pale). Hollow bars indicate changes that were not significant at that level. ± 1 standard deviation for each change is shown as a black vertical line.

are significant, but they differ by small amounts: <10%. Only the change in $f_{\bar{R}0}$ is both significant and large (>10%). Bandwidth and magnitude changes for all resonances and anti-resonances, with the exception of $\bar{R}3$, are significant at the 1% level. Increases in the bandwidth B were also reported by Pham Thi Ngoc and Badin (1994) for the acoustic resonances, with an increase in B_{R1} of 40%–120% (cf. ~90% here) for the French vowel [a] (where f_{R1} is close to that used in this study) and up to 140% for higher resonances.

Trends in the women's data are less clear, particularly for changes in frequency, but appear to follow the same patterns for B and $|Z|$. Variation may be more prominent because only three women took part in the study.

The impedance of the glottis during phonation is largely due to an effective mass of air between the vocal folds: the inertive reactance of this mass would increase with increasing frequency. This would explain why the acoustic resonances (at high frequency) are less affected by the change from closed glottis to phonation than the low-frequency acousto-mechanical resonance. The experiments on the model system show that B_{R1} increases by 15% with open glottis, and by 35% with a rapid airflow of $6 \, \text{L s}^{-1}$ through the glottis, compared to a closed glottis condition. These changes are smaller than those in Fig. 8.

In summary, for the purely acoustic resonances (i.e., Ri and $\bar{R}i$ where $i \geq 1$) phonation increases B_{Ri} and $B_{\bar{R}i}$ while increasing $|Z|_{Ri}$ and decreasing $|Z|_{\bar{R}I}$ to a greater extent than

predicted by a simple model, perhaps due to a combination of the DC flow increasing visco-thermal losses, turbulence in the jet of air from the glottis, and/or transmission of energy through the glottis to the subglottal tract. For the acousto-mechanical resonances, $f_{\bar{R}0}$ increases and B and $|Z|$ of both $\bar{R}0$ and R0 decrease.

2. Including the glottis in the simple vocal tract model

Several of the changes from closed glottis to phonation can be qualitatively explained by introducing a glottal opening to the vocal tract model. During phonation, the glottis has a non-zero average opening, which changes the load impedance Z_L in Eq. (1). The shape of the airway between the vocal folds is complicated (Šidlof et al., 2008) but to retain the simplicity of the model, the glottis is considered as an open cylinder of length l_g and radius r_g. Dalmont et al. (2001) showed empirically that the impedance of such a cylinder can be represented by

$$Z = -j \frac{\rho c}{\pi r_g^2} \tan\left(\Gamma_g (l_g + \delta)\right), \tag{9}$$

where Γ_g depends on r_g according to Eq. (2), and the additional effective length for a flanged opening $\delta = [0.8216 r_g (1 + [(0.77\Gamma_g r_g)^2/1 + 0.77\Gamma_g r_g])^{-1}]$. Figure 9 shows the impedance calculated from Eq. (1) with Z_L from Eq. (9), where Γ includes the increased α determined in Sec. III A 3, in parallel with the low frequency b, L_t, C_t circuit. In this case the values of $r_g = 1 \, \text{mm}$ and $l_g = 10 \, \text{mm}$ are chosen to give approximate agreement between the calculated curve and the male impedance data in Table II.

The inferred values $r_g = 1 \, \text{mm}$ and $l_g = 10 \, \text{mm}$ are reasonable (Šidlof et al., 2008), although much larger maximum glottal areas have been reported during low pitched phonation—equivalent to $r_g = 3.2 \, \text{mm}$ (Hoppe et al., 2003). However, the average r_g over the time window of the
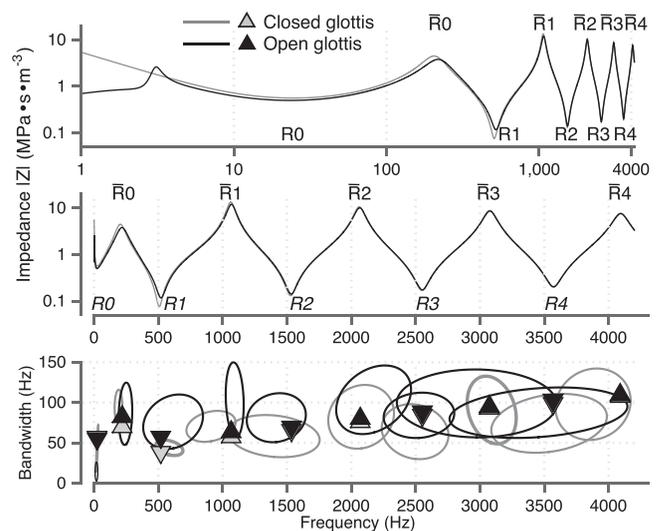

FIG. 9. Two models of the impedance of the vocal tract shown with logarithmic (top) and linear (middle and bottom) frequency axes. Pale: closed glottis as shown in Fig. 5. Dark: $r_g = 1 \, \text{mm}$, $l_g = 10 \, \text{mm}$. B for the closed glottis measurements (pale ellipses), for the model shown above (pale triangles) and for the phonation measurements (dark ellipses) and its model (dark triangles).

J. Acoust. Soc. Am. **139** (5), May 2016

Hanna et al.   2933

measurement is presumably much lower than the maximum value. For an effective aperture $r_g > 1.4$ mm, the subglottal impedance is expected to contribute features to the impedance spectrum measured at the mouth (Fant *et al.*, 1972). Extrema in impedance spectra that might be attributed to subglottal resonances were not observed in the study described here.

In practice, how important are the resonance R0 at ~20 Hz and the anti-resonance R̄0 at ~200 Hz that were measured here (e.g., Figs. 5 and 6)? In vowels, the compliance C of the air in the tract is usually nearly short-circuited by the low value of the radiation impedance at the (open) lips. The 20 Hz resonance would rarely be excited by the signal due to the source at the larynx. However, the pressure transients due to occlusives and perhaps especially plosives excite the tract over a large frequency range extending to very low frequencies. R̄0 does lie within the frequency range of speech and therefore may influence the relative prominence of f_o or its harmonics, particularly when the mouth opening is small. Consequently, the effects of these acousto-mechanical resonances may affect the low frequency spectral components of some consonants.

An interesting observation is that R̄0 may determine the lower limit of f_{R1}. At low frequencies, the inertance of the air at the lip aperture is in parallel with that of the tissue, and both load the compliance of air in the tract. If the tissue inertance was very large, reducing the lip aperture to very small values would decrease f_{R1} to very low values. However, with the finite tissue inertance f_{R1} and therefore $F1$ must be higher than $f_{\bar{R}0}$.

Pham Thi Ngoc and Badin (1994) reported that changes in resonance frequencies from glottis closed to phonation were negligible in French vowels, except in the case of the close (also known as high) vowels [i] and [y], which increase by more than 10% during phonation. They suggest that this effect could be due to differences in tongue position due to absence of feedback from the airflow when the glottis was closed. However, f_{R1} must be greater than $f_{\bar{R}0}$, and the latter here increased by ~15% from closed glottis to phonation (see Fig. 8). This gives another explanation why, for close vowels (such as /i/ and /u/), a low f_{R1} may increase from closed glottis to phonation.

Finally, both mechanical resonances R0 and R̄0 may have an effect on the spectral shape of occlusives. The broadband spectrum of abrupt transients extends to low frequencies and the initial broadband burst of bilabial plosives such as [b] is observed to have a characteristic low frequency prominence and downward spectral tilt (Kewley-Port, 1983).

3. Estimating the impedance seen by the glottis

Since all the measurements presented are made with the subjects' lips sealed around the impedance head, the values of B and α do not include losses associated with the radiation field of the open mouth, and so are not directly comparable with measurements in the literature performed using the voice signal. Transmission loss due to the resistive term of the radiation impedance of an open cylinder causes the magnitudes of its impedance extrema to decrease more rapidly

with frequency than those of a closed cylinder. To approximate this, the simple lossy, flexible cylindrical model discussed above is used to provide an estimate of the impedance as "seen" by the glottis looking toward the open mouth. In this case, the remote termination of the duct is the low radiation impedance at the lips, rather than the high impedance of the glottis.

The vocal tract impedance is calculated from Eq. (1) (where r is the radius of the vocal tract from Table III, l its length from Eq. (4), and Γ includes the increased α) in parallel with the low frequency b, L_t, C_t circuit of Fig. 6. This non-rigid cylinder is terminated by the load impedance of the lips, modeled by a flanged open cylinder [Eq. (9)] with variable radius, and length 10 mm to account for the position of the measurement plane. Note that, over the frequency range of interest, the details of the facial geometry as a baffle are unimportant (Arnela *et al.*, 2013).

This calculation provides a model with an effective bandwidth for comparison with values in the literature. Figure 10 shows the calculated impedance from the glottis with three different effective lip apertures, and thus three different radiation impedances. The calculated bandwidths are superimposed on a range of estimated formant bandwidths based on an algebraic relationship relating Fi and B_{Fi} (Hawks and Miller, 1995) that was based upon previous swept sine excitation measurements performed with a closed glottis (Fant, 1961; Fujimura and Linqvist, 1971) and depends on the fundamental frequency f_o.

In speech the vowel [ɜː] is spoken with an effective radius of the lips that is usually smaller than that of the tract; that case, where the effective lip radius is 70% that of the tract, is shown with the neutral dark curve in the figure. The two extreme cases are also included for reference:
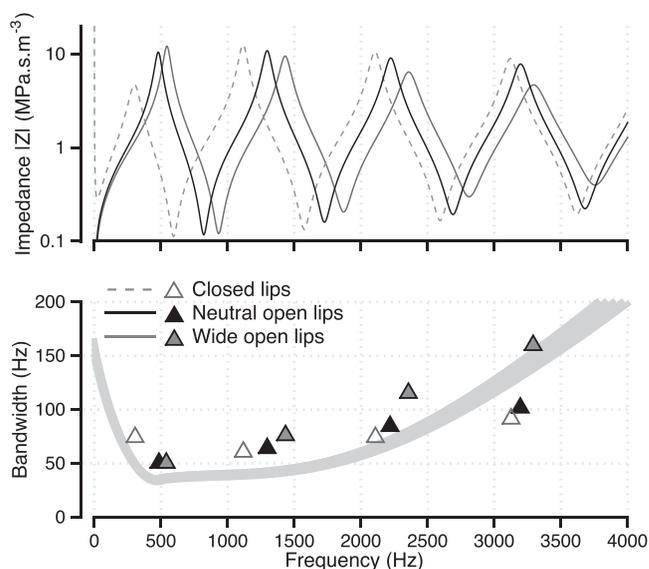

FIG. 10. (Top) The modeled impedance curves as "seen" from the glottis with three different effective lip apertures: the same as the tract (solid pale), a smaller effective radius of 7 mm (solid dark), and closed (pale dashed). (Bottom) The bandwidths calculated from the maxima of these curves are pale filled, dark filled, and dark open, respectively. These are compared to the pale shaded using the formant bandwidth equation of Hawks and Miller (1995) with f_o from 90 to 150 Hz.

closed lips, and an aperture equal to the cross section of the tract.

The closed (or almost closed) lip case also demonstrates that B_{F1} increases as $F1$ decreases. Hawks and Miller found it necessary to describe this behavior by using different sets of parameters for low and high frequency. However, their calculation allows for unrealistically low frequency formants and does not consider how the non-rigidity of the vocal tract influences the lowest achievable $F1$, as shown by Sondhi (1974). For the model in Fig. 10 this lower limit of $F1$ occurs when the lip radius is $<1\%$ of the tract radius. Using the average data from the vocal tract configurations in this study, this lower limit occurs at $\sim$300 Hz for men and $\sim$320 Hz for women. It is possible to further decrease $F1$ by small amounts by lengthening the vocal tract, and for non-cylindrical vocal tracts other changes to the area function will affect the lowest achievable $F1$. However, because the inertance of the air between the lips is in parallel with the impedance of the tissue surrounding the tract, it should not be possible for $F1$ to occur lower than the closed glottis $f_{\bar{R}0}$ reported in Table II.

IV. CONCLUSIONS

A new method yields direct measurements through the lips of the frequency, magnitude, and bandwidths of vocal tract resonances Ri and anti-resonances $\bar{R}$i, both with the glottis closed and during phonation, over a range of 9 octaves and 80 dB dynamic range.

For an articulation approximating the neutral vowel [ɜː] and with glottis closed the impedance magnitudes of purely acoustic vocal tract resonances were around 0.1 MPa s m^{-3} and anti-resonances 10 MPa s m^{-3}. In the absence of a radiation loss from the mouth and with the glottis closed, the bandwidths of the resonances up to 4 kHz increase with frequency from approximately 50 to 90 Hz for men and 70 to 90 Hz for women. Comparison of the measured data with a simple one-dimensional rigid tube gives fair agreement between 300 and 4200 Hz if the attenuation coefficient α is 5 times higher than that calculated with the typical visco-thermal wall losses of a dry, rigid cylinder.

Phonation further increases the bandwidths due to DC airflow and the lossy inertive acoustic load as seen by the vocal tract, but does not significantly change the resonance frequencies. Resonances of the subglottal tract were not observed in the low-pitched, modal phonation performed in this study due to the small average effective glottal aperture.

At lower frequencies, the finite mass and non-rigidity of the vocal tract walls strongly affect the acoustics of the tract causing two acousto-mechanical resonances. A low Q impedance minimum due to the mechanical resonance R0 of the tissue walls of the vocal tract was observed at $\sim$20 Hz, and an impedance maximum $\bar{R}0$ at $\sim$200 Hz due to a resonance involving the tissue mass and the compliance of the air inside the tract. The effective vibrating mass of the vocal tract walls can be considered as a compact mass of 0.17 and 0.15 kg, with a spring constant of 4.3 and 2.1 kN m^{-1}, for men and women, respectively.

The parameters estimated in the study can be incorporated into models of the vocal tract for voice synthesis, so that they may yield more representative transfer functions, formant frequencies, magnitudes, and bandwidths, especially for close vowels (such as /i/ and /u/). Precise measurements of the bandwidths of vocal tract resonances may have the potential to aid clinical diagnosis of obstructive sleep apnea as suggested by Robb et al. (1997).

[1]The terminology of the ANSI (2004) is followed here: a formant is a maximum in the spectral envelope of the sound. See also Titze et al. (2015). (Some speech researchers use formant to mean the resonance that gives rise to the spectral maximum.)

ANSI (**2004**). ANSI S1.1-1994, *American National Standard Acoustical Terminology* (Acoustical Society of America, Melville, NY).

Arnela, M., Guasch, O., and Alías, F. (**2013**). "Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations," J. Acoust. Soc. Am. **134**, 2946–2954.

Boutin, H., Fletcher, N., Smith, J., and Wolfe, J. (**2013**). "Lip motion, the playing frequency of the trombone and the upstream and downstream impedances," in *Proceedings of the Stockholm Music Acoustics Conference*, Stockholm, Sweden, pp. 483–489.

Childers, D. G., and Wu, K. (**1991**). "Gender recognition from speech. Part II: Fine analysis," J. Acoust. Soc. Am. **90**(4), 1841–1856.

Coffin, B. (**1989**). *Historical Vocal Pedagogy Classics* (Scarecrow Press, Maryland), pp. 189–200.

Coltman, J. W. (**2003**). "Acoustical losses in wet instrument bores," J. Acoust. Soc. Am. **114**, 1221.

Dalmont, J. P., Nederveen, C., and Joly, N. (**2001**). "Radiation impedance of tubes with different flanges: Numerical and experimental investigations," J. Sound Vib. **244**, 505–534.

Dickens, P., Smith, J., and Wolfe, J. (**2007**). "Improved precision in measurements of acoustic impedance spectra using resonance-free calibration loads and controlled error distribution," J. Acoust. Soc. Am. **121**(3), 1471–1481.

Dunn, H. (**1961**). "Methods of measuring vowel formant bandwidths," J. Acoust. Soc. Am. **33**, 1737–1746.

Epps, J., Smith, J., and Wolfe, J. (**1997**). "A novel instrument to measure acoustic resonances of the vocal tract during phonation," Meas. Sci. Technol. **8**(10), 1112–1121.

Fant, G. (**1961**). "The acoustics of speech," in *Proceedings of the International Conference on Acoustics*, Stuttgart, Germany, pp. 188–201.

Fant, G. (**1970**). *Acoustic Theory of Speech Production with Calculations Based on X-ray Studies of Russian Articulations* (Mouton, Hague, The Netherlands), pp. 15–26.

Fant, G. (**1972**). "Vocal tract wall effects, losses, and resonance bandwidths," Quarterly Progress Status Report, Vol. 2(3), Department of Speech, Music and Hearing, KTH, Stockholm, pp. 28–52.

Fant, G., Ishizaka, K., Lindqvist, J., and Sundberg, J. (**1972**). "Subglottal formants," Quarterly Progress Status Report, Vol. 13(1), Department of Speech, Music and Hearing, KTH, Stockholm, pp. 1–12.

Fant, G., Nord, L., and Branderud, P. (**1976**). "A note on the vocal tract wall impedance," Quarterly Progress Status Report, Vol. 17(4), Department of Speech, Music and Hearing, KTH, Stockholm, pp. 13–20.

J. Acoust. Soc. Am. **139** (5), May 2016

Hanna et al.   2935

Fant, G., and Sonesson, B. (**1964**). "Speech at high ambient air-pressure," Quarterly Progress Status Report, Vol. 5(2), Department of Speech, Music and Hearing, KTH, Stockholm, pp. 9–21.

Fitch, W. T., and Giedd, J. (**1999**). "Morphology and development of the human vocal tract: A study using magnetic resonance imaging," J. Acoust. Soc. Am. **106**, 1511–1522.

Flanagan, J. L. (**1972**). *Speech Analysis; Synthesis and Perception* (Springer-Verlag, Berlin).

Fletcher, N. H., and Rossing, T. D. (**1998**). *The Physics of Musical Instruments*, 2nd ed. (Springer, New York), pp. 193–198.

Fujimura, O., and Lindqvist, J. (**1971**). "Sweep-tone measurements of vocal-tract characteristics," J. Acoust. Soc. Am. **49**, 541–558.

Hanna, N. (**2014**). "Investigations of the acoustics of the vocal tract and vocal folds in vivo, ex vivo and in vitro," UNSW, Australia and Université de Grenoble, France (https://tel.archives-ouvertes.fr/tel-01174056/), pp. 31–66 (Last viewed 5 November 2015).

Hanna, N., Smith, J., and Wolfe, J. (**2012**). "Low frequency response of the vocal tract: Acoustic and mechanical resonances and their losses," in *Proceedings of the Australian Acoustical Society* (T. McGinn, Fremantle, Australia), pp. 317–323.

Hanna, N. N., Smith, J., and Wolfe, J. (**2013**). "Measurements of the aero-acoustic properties of the vocal folds and vocal tract during phonation into controlled acoustic loads," Proc. Meet. Acoust. **19**, 060174.

Hanson, H. M., and Chuang, E. S. (**1999**). "Glottal characteristics of male speakers: Acoustic correlates and comparison with female data," J. Acoust. Soc. Am. **106**, 1064–1077.

Hawks, J. W., and Miller, J. D. (**1995**). "A formant bandwidth estimation procedure for vowel synthesis," J. Acoust. Soc. Am. **97**(2), 1343–1444.

Hoppe, U., Rosanowski, F., Döllinger, M., Lohscheller, J., Schuster, M., and Eysholdt, U. (**2003**). "Glissando: Laryngeal motorics and acoustics," J. Voice **17**(3), 370–376.

House, A. S., and Stevens, K. N. (**1958**). "Estimation of formant bandwidths from measurements of transient response of the vocal tract," J. Speech Hear. Res. **1**, 309–315.

Ishizaka, K., French, J., and Flanagan, J. (**1975**). "Direct determination of vocal tract wall impedance," IEEE Trans. Acoust., Speech, Signal Process. **23**, 370–373.

Kewley-Port, D. (**1983**). "Time-varying features as correlates of place of articulation in stop consonants," J. Acoust. Soc. Am. **73**, 322–335.

Kob, M. (**2002**). *Physical Modeling of the Singing Voice* (Universitätsbibliothek, Rheinisch-westfälische Technische Hochschule Aachen), pp. 57–70; available at http://sylvester.bth.rwth-aachen.de/dissertationen/2002/132/02_132.pdf (Last accessed 3 September 2014).

Langford-Smith, F. (**1963**). *Radiotron Designer's Handbook*, 6th ed. (Amalgamated Wireless Valve Co. Pty. Ltd., Sydney, Australia), pp. 150–151.

Makarov, I., and Sorokin, V. (**2004**). "Resonances of a branched vocal tract with compliant walls," Acoust. Phys. **50**, 323–330.

Meltzner, G. S., Kobler, J. B., and Hillman, R. E. (**2003**). "Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices," J. Acoust. Soc. Am. **114**, 1035–1047.

Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**(2), 175–184.

Pham Thi Ngoc, Y., and Badin, P. (**1994**). "Vocal tract acoustic transfer function measurements: Further developments and applications," J. Phys. IV **C5**, 549–552.

Robb, M., Yates, J., and Morgan, E. (B.) (**1997**). "Vocal tract resonance characteristics of adults with obstructive sleep apnea," Acta oto-laryngo-logica **117**, 760–763.

Rothenberg, M. (**1981**). "An interactive model for the voice source," Quarterly Progress Status Report, Vol. 22(4), Department of Speech, Music and Hearing, KTH, Stockholm, pp. 1–17.

Russell, G. O., and Cotton, J. C. (**1959**). "Causes of good and bad voices" (National Research Foundation, under a subsidy administered by the Carnegie Institute of Washington), pp. 70–71.

Savitzky, A., and Golay, M. J. E. (**1964**). "Smoothing and differentiation of data by simplified least squares procedures," Anal. Chem. **36**(8), 1627–1639.

Šidlof, P., Švec, J. G., Horáček, J., Veselý, J., Klepáček, I., and Havlík, R. (**2008**). "Geometry of human vocal folds and glottal channel for mathematical and biomechanical modeling of voice production," J. Biomech. **41**, 985–995.

Smith, J. (**1995**). "Phasing of harmonic components to optimize measured signal-to-noise ratios of transfer-functions," Meas. Sci. Technol. **6**(9), 1343–1348.

Sondhi, M. M. (**1974**). "Model for wave propagation in a lossy vocal tract," J. Acoust. Soc. Am. **55**, 1070–1075.

Sondhi, M. M. (**1986**). "Resonances of a bent vocal tract," J. Acoust. Soc. Am. **79**(4), 1113–1116.

Story, B. H., Titze, I. R., and Hoffman, E. A. (**1996**). "Vocal tract area functions from magnetic resonance imaging," J. Acoust. Soc. Am. **100**, 537–554.

Summerfield, Q., Foster, J., Tyler, R., and Bailey, P. J. (**1985**). "Influences of formant bandwidth and auditory frequency selectivity on identification of place of articulation in stop consonants," Speech Commun. **4**(1–3), 213–229.

Swerdlin, Y., Smith, J., and Wolfe, J. (**2010**). "The effect of whisper and creak vocal mechanisms on vocal tract resonances," J. Acoust. Soc. Am. **127**, 2590–2598.

Tarnóczy, T. H. (**1962**). "Vowel formant bandwidths and synthetic vowels," J. Acoust. Soc. Am. **34**, 859–860.

Titze, I. R., Baken, R. J., Bozeman, K., Granqvist, S., Henrich, N., Herbst, C. T., Howard, D. M., Hunter, E. J., Kaelin, D., Kent, R., Kreiman, J., Kob, M., Löfqvist, A., McCoy, S., Miller, D. G., Noé, H., Scherer, R. C., Smith, J. R., Story, B. H., Svec, J. G., Ternström, S., and Wolfe, J. (**2015**). "Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization," J. Acoust. Soc. Am. **137**, 3005–3007.

Van den Berg, J. (**1955**). "Transmission of the vocal cavities," J. Acoust. Soc. Am. **27**, 161–168 (1955).

Wakita, H., and Fant, G. (**1978**). "Toward a better vocal tract model," Quarterly Progress Status Report, Vol. 19(1), Department of Speech, Music and Hearing, KTH, Stockholm, pp. 9–29.

Wodicka, G. R., Stevens, K. N., Golub, H. L., and Shannon, D. C. (**1990**). "Spectral characteristics of sound transmission in the human respiratory system," IEEE Trans. Biomed. Eng. **37**(12), 1130–1135.

Wolfe, J., Almeida, A., Chen, J. M., George, D., Hanna, N. N., and Smith, J. (**2013**). "The player-wind instrument interaction," in *Proceedings of the Stockholm Music Acoustics Conference*, Stockholm, Sweden, pp. 323–330.

Wu, L., Xiao, K., Dong, J., Wang, S., and Wan, M. (**2014**). "Measurement of the sound transmission characteristics of normal neck tissue using a reflectionless uniform tube," J. Acoust. Soc. Am. **136**, 350–356.