Low frequency response of the vocal tract: acoustic and mechanical resonances and their losses

Noel Hanna (1,2), John Smith (1) and Joe Wolfe (1)

(1) School of Physics, The University of New South Wales, Sydney, 2052, Australia

(2) Department of Speech and Cognition, GIPSA-lab, Université de Grenoble, France

ABSTRACT

The impedance spectrum of the vocal tract was measured at the lips from 10 Hz to 4.2 kHz using the three-microphone, three-calibration technique. A broadband signal synthesised from sine waves allows high precision measurements irrespective of the fundamental frequency of phonation. From these measurements, the frequencies, magnitudes and bandwidths of the resonances and antiresonances are determined directly. Resonances have impedance magnitudes of 20-100 kPa.s.m^{-3} and antiresonances 2-10 MPa.s.m^{-3}. The bandwidths measured with a closed glottis are typically around 50 Hz for resonances and antiresonances between 400 Hz and 2 kHz and increase slightly outside this range. This is in qualitative agreement with previous measurements estimated from phonations measured outside the mouth. The measured bandwidths are discussed in relation to viscothermal losses at the duct boundary, radiation from the mouth and mechanical losses in the surrounding tissues.

BACKGROUND

The acoustic features that distinguish speech phonemes are largely determined by resonances of the vocal tract. These resonance frequencies, Ri, and how they vary, characterise vowels and are important to many consonants (Fant 1970). The bandwidths of these resonances also provide valuable information for listeners (Summerfield *et al.* 1985, Brown *et al.* 2010), for speaker identification (Childers and Diaz 2000), and provide information about the glottal source (Childers and Wu 1991, Flanagan 1972, Rabiner and Schafer 1978, and Rothenberg 1981). Further, these resonances in singers can be adjusted to boost the radiation of the fundamental and/or one of the harmonics of the voice (resonance tuning; see Joliveau *et al.* 2004a, 2004b, Henrich *et al.* 2011).

At a fundamental level, the magnitude of the bandwidths is vital to speech communication. A large bandwidth entails a low quality factor or Q, so a resonance with a very wide bandwidth would create a weak formant (peak in the spectral envelope): it would give only modest boost to frequencies lying nearby and so would render vowels and some other phonemes difficult or impossible to identify. A resonance with a high Q and a very narrow bandwidth would give a large boost to a voice harmonic that fell close to the resonant frequency. However, the harmonics of the voice would usually not fall within a very narrow bandwidth. For a voice with fundamental frequency 150 Hz, and therefore a harmonic spacing of 150 Hz, the chance that a harmonic would fall in a 10 Hz bandwidth is only about 1/15. We can argue in the reverse direction: that we can recognise phonemes characterised by formants means that the bandwidth must fall in the 'Goldilocks range': large enough to include at least one harmonic in most phonations, but narrow enough to give that harmonic a significant boost. Despite this importance to speech production and perception, we know of no previous direct measurements of the bandwidths of tract resonances during phonation.

Previous studies of the vocal tract resonances have used either the voice sound itself, swept sine excitation at the neck with the glottis closed (Peterson and Barney 1952, Van den Berg 1955, Dunn 1961, Fant 1972, Fujimura and Lindqvist 1971) or a broadband sound or vibration injected outside the neck (Pham Thi Ngoc and Badin 1994) or at the open mouth (Epps *et al.* 1997).

Using the voice restricts the frequency resolution to a value comparable with the fundamental frequency and makes it difficult to estimate bandwidths. Measurements made with the sound source at the neck are subject to an unknown transfer function. As for the broadband technique at the mouth, the low magnitude of the radiation impedance at the mouth makes it difficult to measure the magnitude of the impedance associated with the resonances and thus makes it difficult to calculate the bandwidths.

The technique used for the measurements in this study is impedance measurement at the mouth, with the lips sealed around the impedance head. It suffers from none of the disadvantages listed above. It does however have the disadvantage that sealing the lips around the impedance head (diameter 26 mm) effectively fixes the position of the first antiresonance. Further, it measures from the mouth end, whereas a measurement from the glottis end would be more immediately useful in many applications.

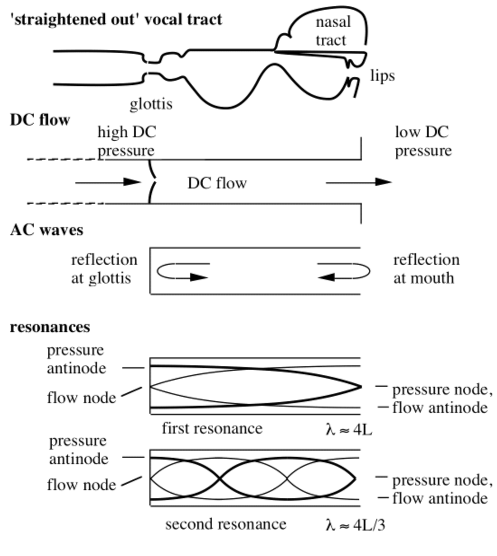

Figure 1. Comparison of a straightened vocal tract with a hard walled pipe (from Wolfe *et al.* 2009).

Figure 1 shows a considerable oversimplification of the tract geometry. For the frequencies of interest, wavelengths are much greater than the transverse dimensions, so a one-dimensional approximation captures much of the physics. In Figure 1 the vocal tract is first shown 'straightened out', which has little effect at the frequencies of interest (Sondhi 1986). Next it is imagined as a cylinder with rigid walls. For the vowel sound in the word 'heard' (denoted [ɜː] in the International Phonetic Alphabet), the acoustic properties over the range 300-3000 Hz are moderately well modelled by such a cylinder. At higher frequencies, the shape of the tract on the cm scale becomes more important, and at lower frequencies, the assumption of rigid wall is less justified (Sondhi 1974). We return to this comparison in the Discussion.

METHOD

Three-microphone three-calibration technique

Precise measurements of vocal tract resonances are made during normal phonation by injecting a broadband sound signal at the lips, using the three-microphone, three-calibration technique developed at UNSW (Dickens *et al.* 2007). The subjects' lips were sealed around the measurement system and the impedance spectrum was measured from 14 Hz to 4.2 kHz using a signal synthesised from sine waves (Smith 1995) spaced at 2.69 Hz (44.1 kHz/2^{14}). One complete cycle of this broadband signal will last some 370 ms. A narrow tube in parallel with the loudspeaker allows a DC airflow so measurements can be made during phonation.

The upper frequency limit of measurements is determined by the smallest spacing between two of the three microphones, which here causes a singularity at 4.3 kHz. The lower frequency limit depends on the loudspeaker, which outputs comparatively little power at low frequencies, and the microphones (B&K 4944A), which have significantly reduced sensitivity below 10 Hz.

The power of the measurement signal is distributed over hundreds of different frequencies, while the power of the voice is concentrated on a relatively small number of harmonics. This limits the possible signal:noise ratio. Thus, while it is possible to cover a very wide frequency range (e.g. 10-4.2 kHz) with a single measurement, even during phonation, using narrower frequency ranges improves the signal/noise ratio and results in better measurements from shorter time periods, particularly when loud phonation is involved.

The presence of the injected low frequency sound can excite mechanical vibration of the tissues surrounding the vocal tract at low frequencies. As we explain below, there is a mechanical resonance at ~20 Hz. Strong excitation of this resonance can perturb normal phonation, adding a strong vibrato that cannot be controlled by the subject, and that can also alter the vocal tract resonances in some cases. For this reason, we made measurements covering four frequency ranges: two to characterise the low frequency mechanical resonances (below 300 Hz), another for the higher frequency acoustic resonances (above 200 Hz) and a fourth covering the whole frequency range.

The lowest frequency feature is the first minimum in impedance, due to the mechanical resonance of the cheeks and the tissue under the jaw. This was studied using a broadband signal over the frequency range of 10-50 Hz and a frequency spacing of 0.34 Hz (44.1 kHz/2^{17}). The first maximum in impedance also involves tissue motion, discussed below. To study both the first maximum and minimum, a range of 14-300 or 400 Hz and resolution of 0.67 Hz (44.1 kHz/2^{16}) was used. Finally, for the acoustic resonances that correspond to the formants of speech (*R*1-*R*5), a frequency range of 200-4200 Hz with a resolution of 2.69 Hz (44.1 kHz/2^{14}) was used.

The subjects (seven male, three female) were asked to find a comfortable position for their mouths ensuring an air-tight seal with their lips around the impedance head. With the mouth open and the tongue low in the mouth the sound corresponded approximately to the vowel [ɜː]. Five of the subjects placed their teeth outside the head during all measurements, while the other five did not. Because the mouth is sealed the frequency of the first resonance is effectively fixed, so by keeping the tongue position constant (tongue low) repeatable measurements that are comparable across subjects were possible. Any effect caused by different positions of the teeth would be evident at higher frequencies that are not important for this study.

During the low frequency measurements (10-50 Hz and 14-400 Hz), four of the male subjects had a small magnet attached to their cheek and/or neck in order to measure the velocity of the tissue by recording the induced EMF in a coil of wire at a fixed distance from the magnet. The voltage-velocity relation for the magnet and coil was calibrated by oscillating the magnet with a shaker and measuring simultaneously the vibration amplitude and the induced EMF in the coil.

RESULTS AND DISCUSSION

Vocal tract model

Before presenting results, we continue with the simple model of Figure 1 to allow qualitative comparison with measurements. First, consider a cylinder with length L = 160 mm and diameter 30 mm, closed at the far (glottis) end, and with rigid walls. The calculated impedance spectrum (magnitude and phase) is shown in Figure 2 (pale dotted line). Maxima occur when $L = \lambda/2$, $2\lambda/2$, $3\lambda/2$ etc, *i.e.* near 1, 2, 3 kHz etc. Minima occur when $L = \lambda/4$, $3\lambda/4$, $5\lambda/4$ etc.

At very low frequencies, the impedance of a closed, entirely rigid cylinder is very large: with DC flow, the pressure would rise until the walls ruptured. The vocal tract is not rigid, of course. To model this in a very simple way, we consider the tract to be an acoustically compact object at low frequencies where its dimensions are << λ, so the pressure is uniform within the tract. The walls can be simply modelled as having a mass m that is accelerated by the (approximately uniform) low frequency pressure in the tract. This mass is supported on tissues that we model as having a spring constant k and a linear loss coefficient R. If the air pressure acts on an area A of tissue, the tissue surrounding the tract has an acoustic inertance $L_t = m/A^2$ and an acoustic compliance $C_t = A^2/k$. The low frequency model of the tract now comprises these elements, plus the compliance C of the tract, treated as a compact object at these frequencies. The equivalent circuit is sketched in the inset to Figure 2.

At these low frequencies, the air in the tract is approximately a compliance in parallel with the tissue, because pressure in the tract causes displacements both by displacing the tissue and compressing the air in the tract. The calculated response of the closed cylinder with walls having parameters m, k and R as described above is also shown in Figure 2 (pale dashed line). There is a new low frequency minimum, corresponding to the series resonance of L_t and C_t, and above that, a new low frequency maximum, corresponding to the parallel resonance of L_t and C.

Low frequency behaviour of the vocal tract

Figure 3 shows measurements of the impedance of the tract made, first with the glottis closed (*cf.* closed tube) and no phonation, then during phonation. Note the qualitative similarity to the dashed line in Figure 2.

The roughly equal spacing of the acoustic maxima and minima are explained by noting that, with the mouth aperture fixed around the measurement head, and the tongue low in the mouth, the configuration is close to that of the neutral vowel [ə:], which is known to have roughly equally spaced formants and so is the vowel that best approximates a simple cylinder. Departures from equal spacing at very high frequencies are probably due to the non-uniform cross-sectional area of the vocal tract and departures from cylindrical shape on the cm scale.

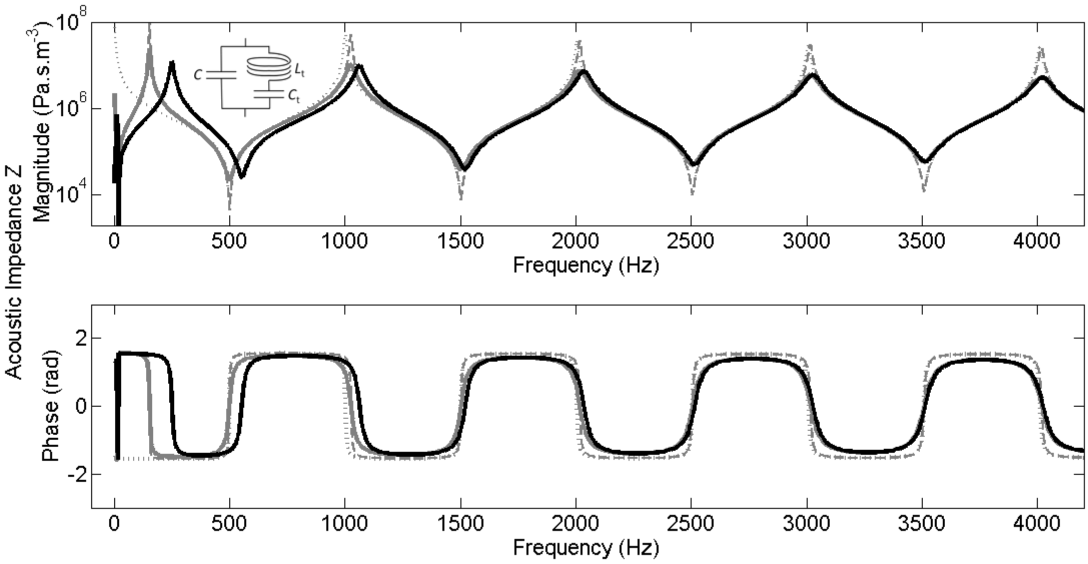

Figure 2. Theoretical impedance (magnitude and phase) of the simple models. Dotted pale line: a rigid-walled, closed cylindrical tube. Pale dashed line: cylindrical tube with finite wall mass and compliance. Pale continuous: a cylindrical tube with finite wall mass and compliance and a loss factor five times greater than the visco-thermal losses for a rigid tube. Dark continuous: as previously, but with the glottis represented by an opening area 32 mm^2, length 14 mm in accordance with the values from (Hoppe *et al.* 2003). The inset shows an electrical equivalent circuit of the low frequency model where the air in the tract is a compact compliance *C* and the tissue surrounding the tract has an acoustic inertance $L_t = m/A^2$ and an acoustic compliance $C_t = A^2/k$.

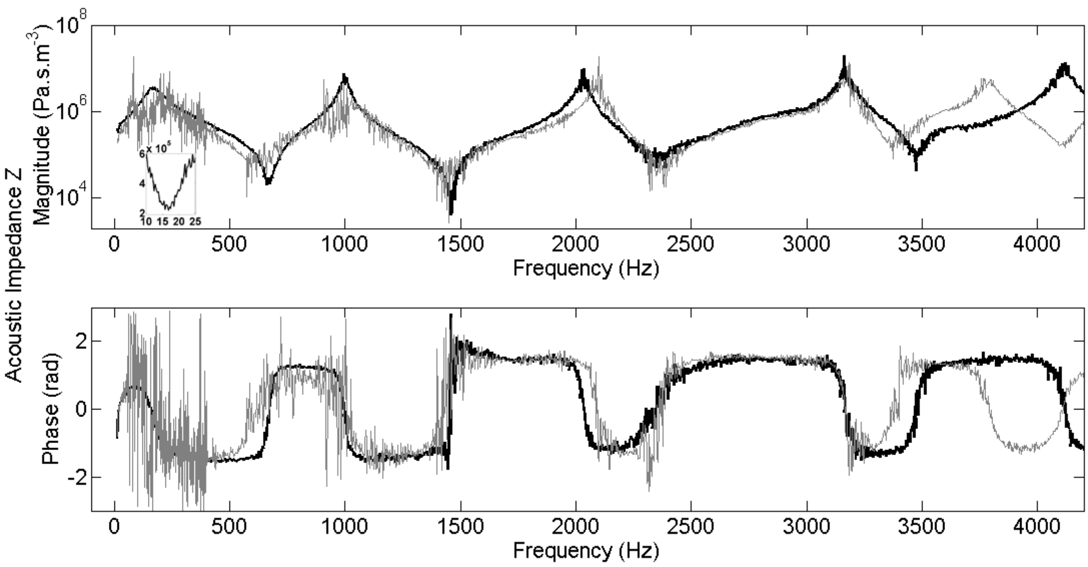

Figure 3. Measured impedance (magnitude and phase) taken from a single cycle of the broadband signal for one subject with closed glottis (dark), and phonating in mechanism 1 (pale). The data from two or more such cycles measured during the same phonation are used to produce the mean and standard deviations shown in Figure 4 and Figure 5. The inset on the upper plot shows a linear plot of the frequency range around the first minimum.

The solid lines in Figure 2 have a larger loss factor α, as described in the Discussion, to produce a curve whose extrema have magnitudes comparable with those in Figure 3.

The plots in Figure 3 are taken from measurements of one subject during only a single cycle of the broadband signal lasting 0.37 s. The pale line measurement displays noise produced by the air flow associated with phonation. Subsequent data analysis is performed on the mean of several such measurements during the same phonation gesture.

As expected, the finite mass and rigidity of the tissues surrounding the tract give rise to the extra minimum and maximum at low frequency: the first minimum is 300-500 kPa.s.m^{-3} at ~20 Hz due to the series resonance of the tissue oscillating on its own elasticity; the first maximum is 3-4 MPa.s.m^{-3} at ~200 Hz, which we attribute to the same tissue oscillating on the 'spring' of the enclosed air.

In all our measurements, the ratio of the frequencies of the first minimum and maximum was about one order of magnitude. This allows very simple modelling to estimate the tissue inertance and compliance L_t and C_t. First, we note that the compact compliance of the air in the tract, C, is determined by the volume of the vocal tract. C is approximately an open circuit at low frequency to a reasonable approximation, so the first minimum occurs at the series resonance

$$f_{0\min} \approx 20Hz \sim \frac{1}{2\pi\sqrt{L_t C_t}} \qquad (1)$$

At high frequency, C_t is approximately a short circuit, so the first maximum occurs at the parallel resonance

$$f_{0\max} \approx 200Hz \approx \frac{1}{2\pi\sqrt{L_t C}} \qquad (2)$$

Taking a volume of V~1x10^{-4} m^3 as an approximate value for the tract, these two equations yield L_t and C_t from any closed glottis measurement.

The low frequency compliance of the air in the vocal tract can be estimated from standard atmospheric conditions, where γ is the adiabatic constant and P_A is atmospheric pressure, so

$$C = V/\gamma P_A \sim 7\text{x}10^{-10}\ \text{m}^3\text{Pa}^{-1} \qquad (3)$$

Substituting C into Equation (2) gives a tissue inertance L_t of ~ 900 kg.m^{-4}. This can be used to give an estimate of the thickess of tissue, w, involved in the vibration. Assuming that the vocal tract is cylindrical we can take its inner surface area A as of the order 0.01 m^2. This is very roughly the area of the tissues (cheeks and under the jaw) that is observed to vibrate. The density ρ of the tissue surrounding the tract is about 10^3 kg.m^{-3}, then Equation (4) gives a plausible tissue thickness w of ~0.01 m.

$$L_t = \frac{m}{A^2} = \frac{\rho V}{A^2} = \frac{\rho w}{A} \qquad (4)$$

The compliance of the tissue can now be determined by substituting L_t into equation (1), giving C_t ~ 7x10^{-8} m^{3}Pa^{-1}. The spring constant of the tissue k can be determined by considering

$$C_t = \frac{dV}{dP} = \frac{A\,dx}{dF/A} = \frac{A^2}{k} \qquad (5)$$

Again assuming an area A = 0.01 m^2, the calculated spring constant is k ~ 1000 Nm^{-1}, which appears reasonable for biological tissue. A summary of these values is given in Table 1.

The amplitudes and distributions of vibrations in the cheeks and neck are not reported here but are consistent with the pattern mapped out by Fant *et al.* (1976).

Bandwidth and Q factor

In order to estimate the bandwidth and quality factor Q of the resonances, a Savitzky–Golay smoothing filter (Savitzky and Golay 1964) was used to identify the maxima and minima of the magnitude of acoustic impedance. A parabola was then fitted locally to the data near these maxima and minima. The bandwidth was determined as the frequency range below 1/√2 times the maximum impedance for antiresonances or √2 times the minimum impedance for resonances.

As it is difficult for subjects to maintain exactly the same gesture over several seconds without moving, averages were taken over two or more cycles, during which the impedance magnitude curves were similar. The mean bandwidths calculated from such averages for a single subject are shown in Figure 4.

For this and other measurements with closed glottis, the bandwidths for the acoustic resonances (those above 250 Hz) are approximately 50 Hz, with a slight increase at higher frequencies. Such an increase is expected due to the increase with frequency of the visco-thermal losses at the boundary layer with the walls of a tube. A similar weak dependance on frequency can also be seen in the decreasing magnitude and increasing bandwidths of the first two theoretical curves in Figure 2.

	Tract	Tissue		
Acoustic	C = 7x10^{-10} m^{3}Pa^{-1}	L_t = 900 kg.m^{-4}	C_t = 7x10^{-8} m^{3}Pa^{-1}	(Fitted to Data)
Geometric	V = 10^{-4} m^3	A = 0.01 m^2		(Assumed)
		w = 0.01 m	k = 1000 Nm^{-1}	(Calculated)

Table 1. Low frequency model resonance parameters

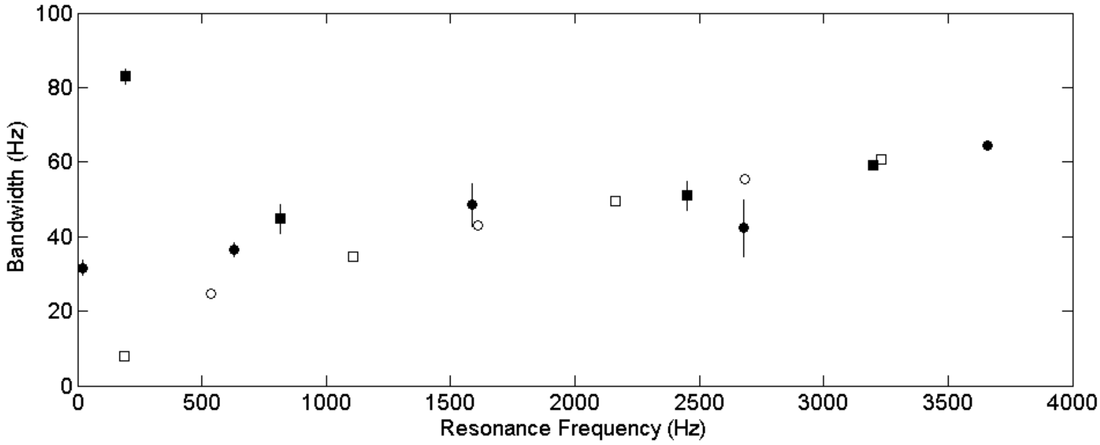

Figure 4. Mean measured bandwidth as a function of resonance frequency for one subject with glottis closed. Closed circles indicate impedance minima, closed squares indicate impedance maxima. Data are for multiple measurements and error bars show ±1 standard deviation. The open shapes show the bandwidths of the cylindrical tube with finite wall mass and compliance and a loss factor α five times greater than the visco-thermal losses for a rigid tube.

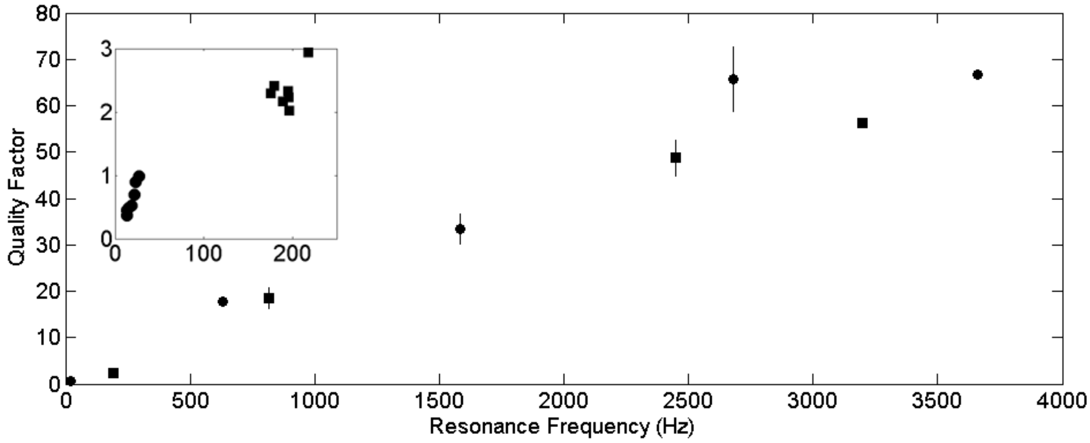

Figure 5. Mean Q factor as a function of resonance frequency measured for one subject with glottis closed. Circles indicate impedance minima, squares indicate impedance maxima, error bars show ±1 standard deviation. Inset: an expansion for the first impedance minimum and maximum.

Figure 5 shows the Q factor (ratio of frequency to bandwidth) as a function of resonance frequency for the same measurements as Figure 4. The Q factor for the acoustic bandwidths shows a dependance on the frequency consistent with losses that increase with increasing frequency.

Because these measurements were made with the impedance head sealed by the subjects' lips, the bandwidths shown do not include radiation loss from the open mouth. When the mouth is opened, the boundary condition at the mouth—the radiation impedance at the mouth—is largely inertive, but it also includes a real, resistive term. Because of transmission and hence loss of energy at the open end, the impedance spectrum of an open cylinder shows impedance extrema whose magnitudes decrease more strongly with frequency than do those of a closed cylinder. So the next step is to use the losses obtained from these experiments to estimate the bandwidths that would apply to the vocal tract with the mouth open.

In this preliminary study, the method used is simply to find the value of the (complex) visco-thermal loss factor α that, when applied to a cylindrical closed tube, gives an impedance spectrum with extrema values similar to those measured here. An increase in α by a factor of five, shown in Figure 4, was required to produce the solid line plots in Figure 2, whose extrema are similar to those of the experimental measurement in Figure 3. To model the input impedance of the tract as 'seen' from the glottis, we use this factor of α to calculate the input impedance at one end of a cylinder that is open and flanged at the far end. This gives an estimation of the impedance at the glottis, which is shown in Figure 6. The sound source at the glottis does not have a cross section equal to that of the tract, however. So Figure 6 shows the impedance of the open cylinder, with the higher value of α, as would be measured through a glottal aperture of radius 3.2 mm (Hoppe *et al.* 2003), including its flanged end effect. Larger radii are also shown to illustrate the effect of widening the glottis. The first, low frequency minimum and maximum are not shown here because, viewed from the glottis end, the air in the tract cannot be treated as a compact object. The contribution of finite tissue compliance to the transfer functions with the open mouth will be reported in a later study. Note that no supra-glottal constriction is included here: such a constriction can have a considerable effect on the impedance at the glottis end of the tract, particularly on the higher formants.

The bandwidths of the models used plotted in the upper part of Figure 6 are shown in the lower part of the figure. The measurements in Figure 3 show that phonation further increases the bandwidths, perhaps because DC flow during phonation increases viscothermal losses, and possibly because of turbulence in the jet emanating from the glottis.

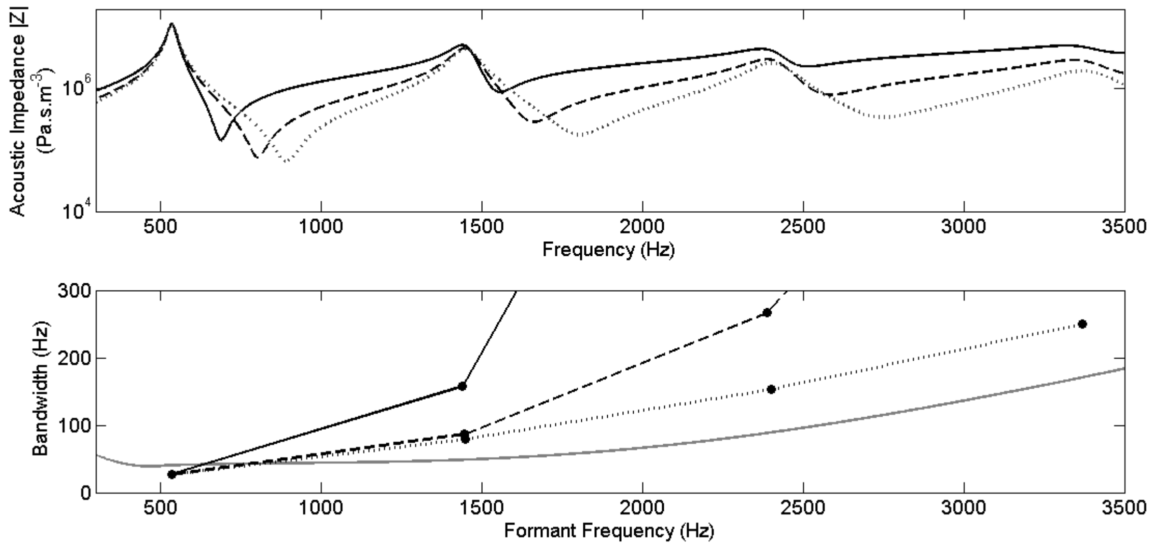

Figure 6. Calculated impedance spectra of a simplified vocal tract as seen at the open glottis, using the loss factor α estimated from Figure 4. The glottal aperture is modelled as the end effect of a tube of radius of 3.2 mm (solid), and 9.6 mm (dashed), and with the same radius as the vocal tract (dotted). The bandwidths of the maxima of these three models are shown in the lower plot. Note that with the smaller glottis openings some of the higher resonances have bandwidths that are undetermined. The smooth line shows an estimation of formant bandwidths for a phonation frequency of 150 Hz using the equation from Hawks and Miller (1995) (this curve would have large error bars not shown here).

The next step is to compare these measurements and calculations of bandwidth with the bandwidths of speech formants, the broad maxima in the spectral envelope of phonated speech. Formants cannot be precisely measured, because the frequency domain is only sampled at multiples of the fundamental frequency f_0. Hawks and Miller (1995) analysed previous measurements made with swept sine excitation and found a strong dependence on f_0, which may be related to this sampling. They published a formant bandwidth estimation curve, shown as an example in the lower part of Figure 6.

The preliminary results reported here will soon be augmented by a much larger data set and a more detailed analysis, which will be presented in the oral version of this paper. These will allow us to relate the bandwidths of formants to measured acoustical losses of the tract. The importance of such a comparison is considerable, because the value of the losses have important implications for perception of speech: very high losses would give broad resonances and broad but weak formants. Low losses would give narrow, strong resonances but, if the bandwidths were too narrow, few or no harmonics would be boosted by them. For example, at a steady phonation frequency of 150 Hz, a bandwidth of 50 Hz implies that there is a one-in-three chance that the harmonics of the voice will experience a substantial amplitude boost from a resonance of the vocal tract. In speech, where the fundamental frequency of the voice often changes during the course of a phoneme, the chances of experiencing a boost are somewhat greater.

The low frequency measurements provide data relating to the mechanical properties of the tissue and its interaction with the tract acoustics. The lowest mechanical resonance at ~20 Hz has a very low Q factor ~1, suggesting that the vibration of the tissue acts as a mass on a very lossy spring. The resonance maximum at ~200 Hz has a Q factor that is still fairly low: ~2. Although the 'spring' involved is that of the air, the motion of the tissue is also damped by the losses in deforming the tissue itself.

CONCLUSIONS

The frequency, magnitude and bandwidths of vocal tract resonances were measured directly both with the glottis closed and during phonation.

For an articulation approximating the vowel [ɔː], the impedance magnitudes of purely acoustic resonances were 20-100 kPa.s.m^{-3} and antiresonances 2-10 MPa.s.m^{-3}. The bandwidths of these resonances with closed glottis and in the absence of radiation loss from the mouth were approximately 50 Hz. They depend only weakly on the resonance frequency.

The simple 1-dimensional vocal tract model approximates the measured data above 250 Hz if the losses are set to be several times higher than the visco-thermal wall losses expected for a rigid tube.

At lower frequencies, the vocal tract walls are not rigid and affect strongly the acoustics of the tract. A low Q factor impedance maximum due to the mechanical resonance of the tissue walls of the vocal tract was observed at ~200 Hz. An impedance minimum at ~20 Hz (mouth closed) is due to a resonance of tissue mass and the compliance of the air inside the tract.

ACKNOWLEDGEMENTS

We thank our volunteer subjects and the Australian Research Council for their support. The presentation of this paper was supported by a travel award from the New South Wales division of the Australian Acoustical Society.

REFERENCES

Brown, NJ, Xuan, W, Salome, CM, Berend, N, Hunter, ML, Musk, AW, James, AL & King, GG 2010, 'Reference equations for respiratory system resistance and reactance in adults', *Respiratory physiology & neurobiology,* 172(3): 162-168.

Childers, DG & J. A. Diaz 2000, 'Speech processing and synthesis toolboxes', *Journal of the Acoustical Society of America,* 108: 1975.

Childers, DG & Wu, K 1991, 'Gender recognition from speech. Part II: Fine analysis', *Journal of the Acoustical Society of America*, 90(4): 1841-1856.

Dickens, P, Smith, JR, & Wolfe, J 2007, 'Improved precision in measurements of acoustic impedance spectra using resonance-free calibration loads and controlled error distribution', *Journal of the Acoustical Society of America*, 121(3): 1471-1481.

Dunn, H 1961, 'Methods of measuring vowel formant bandwidths', *Journal of the Acoustical Society of America,* 33: 1737-1746.

Epps, J, Smith JR & Wolfe, J 1997, 'A novel instrument to measure acoustic resonances of the vocal tract during phonation', *Measurement Science & Technology*, 8(10): 1112-1121.

Fant, G 1970, *Acoustic theory of speech production with calculations based on X-ray studies of Russian articulations* Mouton De Gruyter.

Fant, G 1972, 'Vocal tract wall effects, losses, and resonance bandwidths', *Speech Transmission Laboratory Quarterly progress and status report,* 2(3): 28-52.

Fant, G, Nord, L & Branderud, P 1976, 'A note on the vocal tract wall impedance', *Speech Transmission Laboratory Quarterly progress and status report,* 4(1976): 13-20.

Flanagan, JL 1972, *Speech analysis; synthesis and perception* Berlin, New York, Springer-Verlag.

Fujimura, O & Lindqvist, J 1971, 'Sweep-Tone Measurements of Vocal-Tract Characteristics', *Journal of the Acoustical Society of America,* 49: 54 -558.

Hawks, JW & Miller, JD 1995, 'A formant bandwidth estimation procedure for vowel synthesis', *Journal of the Acoustical Society of America*, 97(2): 1343-1444.

Henrich, N, Smith, JR & Wolfe, J 2011, 'Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones', *Journal of the Acoustical Society of America,* 129: 1024-1035.

Hoppe, U, Rosanowski, F, Döllinger, M, Lohscheller, J, Schuster, M & Eysholdt, U 2003, 'Glissando: laryngeal motorics and acoustics', *Journal of Voice*, 17(3): 370-376.

Joliveau, E, Smith, J & Wolfe, J 2004a, 'Tuning of vocal tract resonance by sopranos', *Nature,* 427; 116.

Joliveau, E, Smith, J & Wolfe, J 2004b, 'Vocal tract resonances in singing: The soprano voice', *Journal of the Acoustical Society of America*, 116(4): 2434-2439.

Peterson, GE & Barney, HL 1952, 'Control methods used in a study of the vowels', *Journal of the Acoustical Society of America*, 24(2): 175-184.

Pham Thi Ngoc, Y & Badin, P 1994, 'Vocal tract acoustic transfer function measurements : further developments and applications', *J. Phys. IV*, C5: 549–552.

Rabiner, LR & Schafer, RW 1978, *Digital processing of speech signals* Prentice-hall Englewood Cliffs, NJ.

Rothenberg, M 1981, 'An interactive model for the voice source', *Speech Transmission Laboratory Quarterly progress and status report*: 1-17.

Savitzky, A & Golay MJE 1964, 'Smoothing and differentiation of data by simplified least squares procedures', *Analytical chemistry,* 36(8): 1627-1639.

Smith, JR 1995, 'Phasing of Harmonic Components to Optimize Measured Signal-to-Noise Ratios of Transfer-Functions', *Measurement Science & Technology*, 6(9): 1343-1348.

Sondhi, MM 1974, 'Model for wave propagation in a lossy vocal tract', *Journal of the Acoustical Society of America,* 55: 1070-1075.

Sondhi, MM 1986, 'Resonances of a bent vocal tract', *Journal of the Acoustical Society of America,* 79(4): 1113-1116.

Summerfield, Q, Foster, J, Tyler, R & Bailey, PJ 1985, 'Influences of formant bandwidth and auditory frequency selectivity on identification of place of articulation in stop consonants', *Speech Communication*, 4(1-3): 213-229.

Van den Berg, J 1955, 'Transmission of the vocal cavities', *Journal of the Acoustical Society of America,* 27: 161-168.

Wolfe, J, Garnier, M & Smith, J 2009, 'Vocal tract resonances in speech, singing, and playing musical instruments', *HFSP Journal,* 3(1): 6-23.