Title: Acoustic analysis of primate air sacs and their effect on vocalization.

Author:

Bart de Boer

Universiteit van Amsterdam

Spuistraat 210

1012 VT Amsterdam

b.g.deboer@uva.nl

Running title: Acoustic analysis of primate air sacs.

Date: 1 October 2009

Abstract:

This paper presents an analysis of the acoustic impedance of primate air sacs and their interaction with the vocal tract. A lumped element model is derived and it is found that the inertance of the neck and the volume of the air sac are relevant, as well as the mass and stiffness of the walls (depending on the tissue). It is also shown that at low frequencies, radiation from the air sac can be non-negligible, even if the mouth is open. It is furthermore shown that an air sac can add one or two low resonances to the resonances of the oral tract, and that it shifts up the oral tract's resonances below approximately 2000 Hz, and shifts them closer together. The theory was verified by acoustic measurements and applied to the red howler monkey (*Alouatta seniculus*) and the siamang (*Symphalangus syndactylus*). The theory describes the physical models and the siamang calls correctly, but appears incomplete for the howler monkey vocalizations. The relation between air sacs and the evolution of speech is discussed briefly, and it is proposed that an air sac would reduce the ability to produce distinctive speech, but would enhance the impression of size of the vocalizer.

PACS-numbers: 43.80 Ka, 43.70 Bk, 43.70 Aj, 43.20 Ks

I. INTRODUCTION

Air sacs are a feature of many primate (and indeed many mammal) vocal tracts. They are air-filled cavities that are connected to the vocal tract. Hewitt et al. (2002) and Riede et al. (2008) describe four main types of primate air sacs. The first type consists of lateral ventricular air sacs, which either exist in pairs or merge ventrally into a single air sac. They connect laterally to the larynx through the ventricles (which are located above the vocal folds) and extend ventrally or laterally. The second type consists of subhyoidal air sacs, connected to the larynx just above the vocal folds, and extending under the hyoid. Subhyoidal air sacs can either be contained completely in an extension of the hyoid (the bulla), or be only partly contained in the hyoid. The third type consists of infraglottal air sacs, connected to the larynx at the level of or below the vocal folds, and extending ventrally. The fourth type consists of dorsal air sacs, which connect between the larynx and the trachea (below the vocal folds) and extend dorsally. Frey et al. (2007), who discuss air sacs in other mammals than primates as well, distinguish a fifth type of air sac, the ventral laryngeal air sac, which exists inside a thyroid bulla, connects to the vocal tract above the vocal folds, and extends ventrally in the laryngeal cavity. Most air sac types have soft walls (and therefore can change volume), but the subhyoidal air sacs that are contained completely in the hyoid bulla have hard walls and fixed volume. For more details, the reader is referred to Hewitt et al. (2002), Riede et al. (2008), Frey et al. (2007) and references therein.

This paper investigates the effect of air sacs on the resonance properties of the primate vocal tract, and is therefore concerned primarily with air sacs that connect to the vocal tract *above* the glottis: the lateral ventricular and subhyoidal air sacs. These are illustrated

in Fig. 1. Subhyoidal air sacs are found among others in howler monkeys (genus *Alouatta*) which are known for their extremely loud and low frequency calls. Lateral ventricular air sacs are found in many species of primates, but most relevantly in apes. All apes have such air sacs: chimpanzees (e.g. Gautier, 1971; Nishimura et al., 2007), orangutans (e. g. Avril, 1963; Brandes, 1932; Camper, 1779) and gorillas (e. g. Kleinschmidt, 1938) while Avril (1963) presents a comparison. However, in humans only the ventricle remains as a vestige of the air sac. There is also some evidence, in the shape of fossil hyoid bones that *Homo neanderthalensis* (Arensburg et al., 1989) and *Homo heidelbergensis* (Martínez et al., 2008) did not have air sacs, while *Australopithecus afarensis* (Alemseged et al., 2006) did (most likely lateral ventricular sacs, as these are the ones found in chimpanzees and gorillas). It is therefore tempting to speculate that the loss of air sacs in humans has something to do with the evolution of speech (e. g. Fitch, 2000; Nishimura, 2008), and this may partly explain a recent increase of interest in the study of air sacs in primates.

The acoustic function of these air sacs, however, is not clear. Authors in the biological literature write about their function as resonators (e. g. Avril, 1963; Schön, 1971), as impedance matchers (e. g. Fitch & Hauser, 1995), or about their potential to suppress resonances (Haimoff, 1983). There is even controversy about whether air sacs have an acoustic function at all. Hilloowalla and Lass (Hilloowala & Lass, 1978) have surgically removed air sacs in rhesus macaques (*Macaca mulatta*, which has partly subhyoidal air sacs) and did not find a difference in the acoustics of their calls. Gautier (Gautier, 1971) on the other hand has found that puncturing an air sac in Brazza's monkey (*Cercopithecus neglectus*, which has lateral ventricular air sacs) leads to a decrease in

volume of the produced calls. It has also been proposed that air sacs are an accidental byproduct of forced expiration (Brandes, 1932), that they serve for rebreathing air (Negus, 1949), and recently a case has been made that they can serve to prevent hyperventilation (Hewitt et al., 2002). It is likely that air sacs play a range of different roles, and it is unlikely that they do not have any acoustic role whatsoever, especially in species that both have loud calls and large air sacs, such as siamangs (Haimoff, 1981, 1983) and howler monkeys (Schön, 1971; Schön Ybarra, 1986; Whitehead, 1995).

In order to make the discussion of the acoustic role of air sacs more concrete, an accurate theoretical model is needed. The purpose of this paper is to develop such a model. This model will analytically describe the acoustic effects of air sacs connected to a vocal tract. It will make it possible to predict (for low frequencies) the main acoustic effects for any primate air sac on the basis of anatomical measurements, without the need for complex computational simulations. However, the correspondence is by necessity approximate. As a first effort to understand the effect of air sacs, the work is therefore more comparable to early work on side branches (such as reviewed in e. g. Fant, 1960; Flanagan, 1965) rather than more quantitatively precise work that takes anatomical detail into account (e. g. Dang & Honda, 1997; Dang *et al.*, 1994). The theory is also used to develop a computer model for higher frequencies.

The work presented here and elsewhere (de Boer, 2008a, 2008b) is closely related to the work presented by Riede et al. (Riede et al., 2008). However, the aim and scope of the present paper are different. Riede et al.'s paper is mostly experimental in nature, and their computational model serves to validate their experiments. Another difference is that Riede et al.'s paper investigates the interaction between the vocal folds and the vocal tract

including the air sac. Although this interaction is likely to be extremely important in ape vocalizations (e. g. Titze, 2008) it falls outside the scope of the present paper.

The paper consists of two parts. In the first part, the theoretical model is developed, it is shown which properties of the air sac are important and an analysis is made of how an air sac influences the resonances of the vocal tract to which it is connected. In the second part, the theory is put to the test by comparing its predictions with measurements on a physical model. The theory is also applied to animal data. Calls from two species that are known for their air sacs, siamangs (*Symphalangus syndactylus*) and red howler monkeys (*Alouatta seniculus*) are predicted on the basis of anatomical data, and the predictions are compared with the actual calls.

II. DEVELOPING A MODEL

A. A lumped element model

A lumped element model is a simplified representation of an acoustic system in which parts of the system are approximated by components that each represent a compliance, a resistance or an inertance. In such a system each component represents a spatially extended part of the acoustic system. This means that lumped element models are only good approximations at low frequencies – corresponding to wavelengths that are long in comparison with the dimensions of the modeled parts. Given that ape air sacs have diameters of approximately 10 cm, and that the minimal allowed wavelength must be larger than two times this length, the maximal frequency for which the model developed here is valid is *lower* than $35000/20 \approx 1750$ Hz, and therefore a rounded down limit of 1500 Hz was assumed. However, as this approach simplifies the system to be modeled, analysis of a lumped element model can provide theoretical insights into the system's

behavior. The elements that are relevant for the acoustics of an air sac are the volume of air that resonates in the air sac, the wall that might vibrate and radiate sound, and the tube (the neck) with which the air sac is connected to the rest of the vocal tract.

For the lumped element analysis, it will be assumed that air in the neck moves with uniform speed, that pressure everywhere in the cavity of the sac is the same, that the wall moves with the same phase and amplitude everywhere, and that the air sac radiates sound as a sphere. The factors that are modeled are illustrated in Fig. 2.

1. The neck

At low frequencies, the air in the neck of the air sac acts as a mass that moves when pressure is exerted on it. There will also be losses caused by viscosity of the air. Such a system can be modeled by a (frequency dependent) resistance in series with an inertance. According to Flanagan (1965, section 3.21), the value per unit length of the inertance is given by:

$$L = \frac{\rho_a}{A} \qquad (1)$$

where ρ_a is the density of air, and A is the cross-sectional area of the tube. The total inertance is given by:

$$L_n = \int_0^l \frac{\rho_a}{A(x)} dx \qquad (2)$$

where l is the length of the tube. For a neck with constant cross-sectional area, this reduces to:

$$L_n = \frac{\rho_a l}{A} \qquad (3)$$

The inertance of an air sac neck of any shape can be expressed as that of an equivalent neck with constant-cross sectional area that has a length and area that are chosen to give the same value for inertance. Given the form of the expression for viscous losses, it is most convenient to choose the minimal cross sectional area as the equivalent area, and calculate the equivalent length accordingly. This is because this form maximizes the viscous losses. In section II.A.6 it will be shown that even under these maximizing conditions, viscous losses can be ignored.

The value for the viscous loss per unit length (following Flanagan, 1965, section 3.22) is given by:

$$R = \frac{S}{A^2}\sqrt{\frac{\omega \rho_a \mu}{2}} \qquad (4)$$

where S is the circumference of the neck, ω the angular frequency, and μ is the viscosity coefficient of air. Integrating over the length of the neck gives the total resistance. It turns out that, given the form of equation (4), assuming constant cross-sectional area equal to the minimal area of the original tract, using the equivalent length giving the same inertance as the original tract and using the maximal circumference of the original tract will always result in a tract *overestimating* the losses. In section II.A.6 it will be shown that even this upper bound on the losses can be ignored for primate air sacs.

The values for inertance and loss at low frequencies for the neck of an air sac are therefore given by:

$$L_N = \frac{\rho_a l_e}{A_{\min}} \qquad (5)$$

and

$$R_n = \frac{S_{max} l_e}{{A_{min}}^2} \sqrt{\frac{\omega \rho_a \mu}{2}} \qquad (6)$$

where l_e is the equivalent length, A_{min} is the minimal cross-sectional area of the original neck, and S_{max} is its maximal circumference.

For frequencies that are low enough for the lumped element model to be valid, this analysis also works for necks that contain branches or for air sacs with two necks (such as the air sac of the siamang). In this case, it suffices to add the volume of the two necks, take the minimal area as the equivalent area and calculate the equivalent length by dividing the total volume of the necks by their minimal area.

2. The cavity

The air in the cavity acts as a spring that stores and releases energy when it is compressed and released and can therefore be modeled as a compliance. As compressed air is heated slightly, and as this heat can be transferred to the walls, some losses might occur. This can be modeled by a (frequency dependent) resistance.

The value of the compliance, C_c in acoustic terms is given by:

$$C_c = \frac{V}{\rho_a c^2} \qquad (7)$$

where V is the cavity volume, ρ_a is the density of air and c is the speed of sound (Fletcher et al., 2004).

The value R_c of the resistance due to losses at the wall is adopted from Flanagan (Flanagan, 1965, section 3.24) and is given by:

$$R_c = \frac{\rho_a c^2}{A(\eta - 1)} \sqrt{\frac{2 c_p \rho_a}{\lambda \omega}} \qquad (8)$$

where A is the cavity's area, η is the adiabatic constant, c_p is the specific heat of air at constant pressure, λ is the coefficient of heat conduction. Values of all constants are given in Table I.

3. The wall

In the lumped element model the wall is assumed to vibrate with the same amplitude and phase everywhere. It is also assumed that it behaves as a damped oscillator, which is described by the following equation:

$$p(t)\cdot A = m\frac{dv}{dt} + bv + k\int v dt \qquad (9)$$

where p is the differential pressure, v is the wall's velocity, m its mass, b its damping and k its stiffness. A free-body diagram of an element of the wall is given in Fig. 3. Differential pressure is the difference between the (vibrating) pressure in the air sac and the pressure of the air outside the air sac.

Acoustic impedance is the ratio between pressure p and volume velocity $u = v \cdot A$. Rewriting equation (9) to volume velocity gives:

$$p(t) = \frac{m}{A^2}\frac{du}{dt} + \frac{b}{A^2}u + \frac{k}{A^2}\int u dt \qquad (10)$$

(as wall motion is small with respect to the radius of the air sac, A is assumed to be constant to first order). Now if p and u are harmonic vibrations, with angular frequency ω, they can be written as:

$$p(t) = Pe^{i\omega t}, \quad u(t) = Ue^{i\omega t} \qquad (11)$$

where P and U are the (complex) amplitudes of the vibrations. Substituting this in (10) and dividing out $e^{i\omega t}$ gives:

$$P = \left(\frac{m}{A^2} i\omega + \frac{b}{A^2} + \frac{k}{A^2} \frac{1}{i\omega} \right) U \qquad (12)$$

This is equivalent to an inertance, a resistance and a compliance in series.

The wall's mass can be calculated from its area, its thickness and its density. This results in the following value for wall inertance L_w:

$$L_w = \frac{m}{A^2} = \frac{\rho_w dA}{A^2} = \frac{\rho_w d}{A} \qquad (13)$$

where ρ_w is the density of wall tissue, and d is the wall's thickness.

The wall's stiffness depends on Young's modulus of the tissue. However, as the deformation of the wall is the stretching of a two-dimensional membrane, and Young's modulus is defined for stretching along one axis, Young's modulus E will have to be converted to a two-dimensional equivalent k_2 using Poisson's ratio ν (illustrated in Fig. 4):

$$k_2 = \frac{E}{2(1-\nu)} \qquad (14)$$

This quantity gives a ratio between stretching of the surface and strain as follows:

$$k_2 = \frac{S/d}{\delta A/A} \qquad (15)$$

where S is the surface tension, d the thickness of the wall, and δA the change in area.

Now in order to find a relation between pressure and volume velocity (as is required for acoustic impedance) a relation between surface tension and pressure is needed, as well as a relation between change in area and volume velocity.

The relation between surface tension and pressure is given by the Young-Laplace equation:

$$p = S\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \qquad (16)$$

where p is the differential pressure, and r_1 and r_2 are the principal radii of curvature. For a spherical air sac, where $r_1 = r_2 = r$ this simplifies to:

$$p = \frac{2S}{r} \qquad (17)$$

In order to calculate the relation between change in area δA and volume velocity u, a relation between a radius of the air sac and its surface is needed. For all three dimensional shapes such a relation has the following form:

$$A = Cr^2 \qquad (18)$$

where C is a constant that depends on the object's shape (4π for a sphere, for example). Now for a small change in area δA this becomes:

$$A + \delta A = C\left(r + \delta r\right)^2 \approx Cr^2 + 2Cr\delta r = A + 2Cr\delta r \qquad (19)$$

For small values of δA and δr, such that second order terms in the small quantities can be ignored, the following relation therefore holds:

$$\delta A = 2Cr\delta r = 2A\frac{\delta r}{r} \qquad (20)$$

or:

$$\frac{\delta A}{A} = 2\frac{\delta r}{r} \qquad (21)$$

Finally, the total displacement of air (equal to the integral of volume velocity over time) is equal to A times δr. Substitution in (21) gives:

$$\delta A = 2\frac{\int u dt}{r} \qquad (22)$$

Taking (15), (17) and (22) together and solving for p finally gives:

$$\frac{4dk_2}{r^2 A}\int u dt = p \qquad (23)$$

From this it follows that the compliance of the wall in terms of properties of the tissue is:

$$C_w = \frac{r^2 A}{4dk_2} \qquad (24)$$

Following Fletcher et al. (2004), the quality factor Q is used to model damping. The wall resistance is calculated from the quality factor and the inertance and compliance of the wall as follows:

$$R_w = \frac{1}{Q}\sqrt{\frac{L_w}{C_w}} \qquad (25)$$

4. Radiation

For simplicity's sake it will be assumed that the air sac is a spherical radiator. This is admittedly quite unlike reality, but the relevant quantity is the amount of acoustic energy radiated per unit area, and this is not too different for different radiators (Flanagan, 1965, figure 3.6). The spherical radiator results in the simplest equations. Furthermore, a hemispherical radiator in an infinite plane baffle (which is more like an air sac on an ape's chest) would result in the same radiation impedance per unit area.

The acoustical impedance Z_R of a spherical radiator is (e. g. Blackstock, 2000, section 10.D.2b):

$$Z_r = \frac{\rho_a c}{A}\frac{1}{1+\frac{1}{ikr}} \qquad (26)$$

where k is the wave number $2\pi f/c$. This is equivalent to a resistance:

$$R_r = \frac{\rho_a c}{A} \qquad (27)$$

in parallel with an inertance:

$$L_r = \frac{\rho_a r}{A} \qquad (28)$$

5. The complete circuit

Four different subcircuits determine the behavior of the air sac. The circuit of the neck consists of a resistance that represents viscous losses in series with the inertance describing the mass of air in the neck. This is a series circuit, as the loss is caused by the air flow. For the wall, the form of equation (12) indicates a circuit of elements in series. The circuit describing the behavior of air in the cavity consists of a resistance that represents compressional losses in parallel to the compliance that describes the compression of the air in the air sac. This is a parallel circuit, as losses and compression are due to the same pressure. The radiation circuit is a parallel circuit because of the form of equation (26).

The circuits describing the wall motion and the outside pressure must be in series, as the volume velocity outside of the wall is equal to that inside of the wall. This circuit must be in parallel to that describing the air in the cavity, as both are subject to the same pressure (that at the end of the neck). As the air flow coming out of the neck must be equal to that compressing the air in the sac and moving the walls, the parallel circuit of cavity and wall must be in series with the circuit describing the neck. The complete circuit is illustrated in Fig. 2.

6. Parameter values

For modeling primate air sacs, it is necessary to estimate a number of parameters. Some of these, such as the properties of air, are relatively well known, and need no discussion. Estimates of Young's modulus and Poisson's ratio for tissue can be found in the literature. However, these values depend very much on the kind of tissue involved. For soft tissue (e. g. Hunter & Titze, 2007) a value of around 10 KPa appears to be realistic. Given that soft tissue consists for a large part of fluids, and is therefore flexible but incompressible, Poisson's ratio must be 0.5 or slightly lower. This results in k_2 being approximately equal to, but slightly lower than Young's modulus, and therefore a value of 9 000 can be used. For cartilaginous tissue, Young's modulus is much higher. For human tracheal rings, values between 2.5 and 7.7 MPa are given by (Lambert et al., 1991), while Jin and Lewis (2003) give values between 0.5 and 2 MPa for articular cartilage. Poisson's ratio of cartilage is assumed to be around 0.45 (Jin & Lewis, 2003, for articular cartilage) resulting in values for k_2 in the same range as that of Young's modulus. Here, a value of 0.45 MPa was used (the lowest value in the range, selected because it best illustrates the effects of intermediate elasticity). As for bony tissue, Martin et al. (1998, table 4.1) give a value of 17.4 GPa for Young's modulus of human cortical bone and 0.39 for Poisson's ratio. This results in a value for k_2 around 15 GPa. The values have very different orders of magnitude, and as will be shown below, result in qualitatively different behavior of the air sac.

A density of tissue of 10^3 kg·m^{-3} was assumed. Values for damping are harder to find in the literature, but in terms of the quality factor they range from 0.5 (which appears to be implicit in Maeda's vocal tract model (Maeda, 1982)) to 10 (Fletcher et al., 2004).

Probably for stiffer materials, the quality factor is higher. A quality factor of 1 has been used for soft tissue, assuming a relatively strong damping. For cartilage and bone a quality factor of 10 was used resulting in less damping. Finally, the radius of the air sac was chosen to be 5 cm, its shape to be spherical and its walls to be 0.5 cm thick.

7. Air sac-oral tract interaction

The next step is to analyze the interaction between an oral tract and an air sac. Riede et al. (2008) also investigate this question, but do not appear to take power radiated from the air sac into account in their computational model. Also, in the present paper a more analytical approach is chosen in order to explain the acoustic effects of air sacs.

Three different vocal tracts roughly corresponding to [a], [ə] and [y] were investigated. All were 16 cm long. The tract for [ə] consisted of a single cylinder of 1.5 cm radius. The tracts for [a] and [y] consisted of two cylinders, each of 8 cm length, and with radii of 2 cm and 1 cm. For [a] the wide tube was nearest to the mouth, and for [y] the narrow tube was nearest to the mouth.

The air sac is modeled as an impedance parallel to the oral tract. This is warranted, as the tube connecting the air sac to the supralaryngeal vocal tract often joins the vocal tract right above the glottis (see Fig. 1 and illustrations in e. g. Hewitt et al., 2002; Riede et al., 2008; Schön Ybarra, 1986). The glottal source was assumed to be an ideal volume velocity source (which has infinite impedance). Although this is an idealization, it makes analysis much easier and is generally considered acceptable as a first approximation (Fant, 1960; Flanagan, 1965; Stevens, 1999). A better approximation would have to take into account that for maximum acoustic output efficiency, the glottal source, the

subglottal system and the supraglottal system (including air sacs) must be matched to each other (Titze, 2002, 2008).

8. Higher frequencies

At higher frequencies, the impedance of the walls increases. Therefore, the motion of the walls becomes negligible, and the difference between the different wall types disappears. It also means that radiation from the air sac becomes negligible. At higher frequencies, the air sac can therefore be analyzed as a system of rigid-walled tubes. As pressure and volume velocity can no longer be assumed constant over lengths comparable to the dimensions of the air sac, analysis in terms of a lumped element model is no longer possible. A transmission line model (such as the model used in Riede et al., 2008) can be used to gain qualitative insights, but because the cross-sectional dimensions of an air sac are large compared to the wave lengths involved, the assumption of one-dimensional wave motion is not valid (the limit of 1500 Hz for the lumped element model then applies to the transmission line model) and therefore no quantitative predictions can be made.
To investigate the resonances and anti-resonances at higher frequencies, finite element models were used. An axisymmetrical geometry was assumed, allowing the use of a two-dimensional mesh. All factors that were identified in the lumped element model, except viscous losses, were taken into account. Viscous losses were not modeled, as this would have considerably complicated the finite element model and their effect is small for the frequencies investigated. The finite element model approximated the Helmholtz equation in the specified geometry, implemented the walls as interface boundaries and implemented radiation using a Bayliss-Turkel first order boundary condition (taken from Givoli, 1991). It was implemented in the GMSH/GETDP software environment (Dular et

al., 1998) using the Galerkin method, and was run to convergence (a relative residual value of 10^{-10} was used). For the computational verification with the finite element model presented in section II.B.3 a triangular mesh of 4667 elements was used.

B. Results

1. Dominant elements of the model

Given the complete circuit, it is possible to determine which elements dominate the circuit's behavior. The values of the different elements of the circuit were calculated for frequencies between 100 and 2000 Hz. The results are shown in Fig. 5. Keeping in mind that parallel impedances are dominated by the lowest impedance, and series impedances by the highest impedance, it is clear that a good estimate of the total impedance of an air sac for the relevant frequency range can be obtained by only looking at the inertance of the neck, the compliance of the cavity and the inertance (mass) and the compliance (stiffness) of the walls. For walls of soft tissue, stiffness can be ignored, while for walls of bony tissue, mass can be ignored. For walls of cartilage, both mass and stiffness play a role. In none of the cases is damping crucial for the frequencies of the resonances (but in the case of cartilage it is relevant for the bandwidth of the wall resonance). The case of soft tissue is in complete agreement with what Fletcher et al. (2004) have found when studying the inflated esophagus of the ring dove. All air sac dimensions within the range that is plausible for primates give this same qualitative result.

The behavior of the air sac in the low-frequency approximation is thus found to be equivalent to that of a circuit consisting of an inertance (representing the air in the neck) in series with a parallel circuit made up of a compliance (representing the cavity air) in

parallel with an inertance (representing the wall mass) and a compliance (representing wall stiffness) in series (the elements are indicated in bold in Fig. 2).

However, for low frequencies and soft walls, the amount of sound that is radiated from the air sac is non-negligible with respect to the amount of sound radiated from the mouth, and therefore radiation has to be taken into account in that case as well.

2. Air sac qualitative behavior

The whole impedance Z_{sac} of the circuit consisting of the air sac and the connecting tube is:

$$Z_{sac} = i\frac{1-\omega^2 L_w C_w - \omega^2 C_w L_n - \omega^2 C_c L_n + \omega^4 L_w C_w C_c L_n}{\omega^3 L_w C_w C_c - \omega C_w - \omega C_c} \qquad (29)$$

From this equation it can be derived that the impedance has a pole (maximum) at angular frequency:

$$\omega_p = \sqrt{\frac{C_w + C_c}{L_w C_c C_w}} \qquad (30)$$

which corresponds to the frequency at which the air in the air sac and the walls resonate. Equation (29), also has a pole at frequency 0, which corresponds to the fact that the air sac cannot be inflated without limits.

The air sac impedance has zeroes[1] at two angular frequencies:

$$\omega_z = \sqrt{\frac{L_w C_w + C_w L_n + L_n C_c \pm \sqrt{\left(L_w C_w + C_w L_n + L_n C_c\right)^2 - 4 L_w C_w L_n C_c}}{2 L_w C_w L_n C_c}} \qquad (31)$$

[1] In the context of impedance "a zero" is a frequency at which the impedance is zero. It should not be confused with the *number* zero.

This is a rather complex expression, so it is instructive to first look at the limits of soft and hard walls. For soft-walled tissue (low stiffness, and therefore infinite C_w) equation (31) reduces to:

$$\omega_z = \sqrt{\frac{L_n + L_w}{L_n L_w C_c}} \qquad (32)$$

Under the same conditions, the frequency of the pole becomes:

$$\omega_p = \sqrt{\frac{1}{L_w C_c}} \qquad (33)$$

For hard-walled tissue (C_w approximately zero) equation (31) reduces to:

$$\omega_z = \sqrt{\frac{1}{L_n C_c}} = c\sqrt{\frac{A_{neck}}{V l_e}} \qquad (34)$$

which is equal to the Helmholtz resonance frequency ω_H (as was to be expected). For the hard-walled case, the pole at non-zero frequency disappears. Equations(32), (33) and (34) suggest a relation between the Helmholtz frequency and the frequencies of the poles and zeros of the soft-walled case:

$$\omega_z^2 = \omega_H^2 + \omega_p^2 \qquad (35)$$

which is equivalent to an equation given by Fant (1972) in relation to resonance frequencies of the human vocal tract when taking wall vibration into account.
Similar qualitative observations can be made about the intermediate stiffness case. Comparing the values of (30) and (31), it is found that one of the zeros has a lower frequency than the pole, and the other zero has a higher frequency. A similar comparison shows that the zero at the highest frequency has a frequency that is higher than that of the zero for the soft walled case, given in equation (32). Comparing equations (30) and (33)

and using the fact that all inertances and compliances are positive, it is clear that the pole frequency of the intermediate case is also higher than the pole frequency of the soft-walled case. A similar comparison shows that the zero with the lowest frequency occurs at a frequency that is lower than that of the wall resonance:

$$\omega_w = \sqrt{\frac{1}{L_w C_w}} \qquad (36)$$

and finally that for the lowest-frequency zero the following condition holds:

$$\omega_{z\,\mathrm{low}}^2 < \omega_p^2 - \omega_w^2 \qquad (37)$$

The qualitative behavior of the air sac impedance at low frequencies can be summarized as follows: in the hard-walled case, there is a single zero at the Helmholtz frequency. For the soft-walled case, there is a pole followed by a zero higher than the Helmholtz frequency. For the intermediate case, there is a zero, followed by a pole, followed by a final zero. The pole and the highest-frequency zero are higher than those in the soft-walled case. The lowest-frequency zero is lower than the resonance frequency of the wall.

3. Computational results

For soft, cartilaginous and bony air sacs with dimensions from Table I, the impedance for frequencies ranging from 20 Hz to 4000 Hz were calculated using the lumped element model. Its absolute value is shown in Fig. 6.

The absolute values of the impedances calculated from the finite element model are given in Fig. 7. The peaks and zeros in the range of 20 Hz –500 Hz correspond to the ones in Fig. 6. Also, the impedances are essentially equal to the ones in Fig. 6 up to about 1500 Hz. At higher frequencies extra resonances and anti-resonances appear. It should be noted that there is no significant difference between the different wall types for these

higher resonances. The first anti-resonance/resonance pair (showing up as a peak followed by a valley around 2300Hz) corresponds to the resonance of the cavity. The isolated anti-resonance (the peak at 2700 Hz) corresponds to the first anti-resonance of the neck, and the final resonance/ant-resonance pair (the valley followed by a peak around 3750Hz) corresponds to the second resonance of the cavity. Note that due to the cavity's spherical shape, its resonances do not occur at the integer multiples expected from a straight tube.

Air sacs with the dimensions given in Table I and with the three different wall types (their impedances calculated with the finite element model) were connected to the three different vocal tracts. A comparison of the total radiated power of the tracts with and without air sacs is given in Fig. 8.

The major changes caused by the addition of an air sac to the vocal tract are that new peaks occur around the frequencies of the lower zeroes in the air sac response. The hard-walled and soft-walled sacs both add one low peak. The cartilaginous-walled air sac adds two extra peaks. The frequency of these peaks is relatively constant. Peaks corresponding to the resonance frequencies of the vocal tracts without air sacs can also be found in the power spectra of the tracts with air sacs. However, these frequencies are shifted up and closer together for all tracts investigated here. Exact values for the frequencies of the peaks in the radiated power of the different tracts can be found in Table II.

Finally, peaks and zeros corresponding to those of the zeros in the air sac impedance at higher frequencies can also be found in the power spectrum of the combined tracts. The position of these peaks is relatively insensitive to the shape of the oral tracts. All the observations made here correspond qualitatively with the findings of Riede et al. (2008).

A major difference between the results presented here and the computational model of Riede et al. (2008) is in the importance of power radiated through the air sac wall. It turns out that most of the acoustic energy that is responsible for the low frequency peaks is radiated from the air sac in the case of soft and cartilaginous walls. This is the result of the fact that at low frequencies, the mass of the air sac wall presents less of an obstacle for the acoustic vibrations, as well as from the fact that the air sac has a larger area than the opening of the mouth, and therefore functions as a much more efficient radiator at low frequencies. The ratio of power radiated through the mouth and power radiated from the air sac for the examples investigated here is presented in Fig. 9.

C. Discussion

A theoretical understanding of the observed changes of the vocal tract response can be gained by inspecting the impedances of the vocal tract without an air sac and that of the air sac. As the glottis is assumed to be a perfect volume velocity source (which has infinite impedance), the maximum energy theorem (e. g. Titze, 2002) states that most energy is transferred when the vocal tract also has infinite impedance. When dealing with infinite impedances, it is easier to work with impedance's reciprocal, the *admittance*. Infinite impedance means zero admittance. This is achieved when the sum of the oral tract's admittance and the air sac's admittance is zero. The absence of an air sac is equivalent to zero admittance, and hence the resonances occur at the frequencies where the oral tract admittance is zero, which are just the resonances of an open/closed tube. When the air sac has non-zero size, its impedance (and its admittance) can be calculated by equation (29) (ignoring radiation and losses). Because all losses are ignored, the real part of the admittance is zero. Therefore, only the imaginary part of admittance, called

susceptance, needs to be studied. The susceptances of air sacs with the dimensions from table I are given in Fig. 10. Also shown (grey line) is the negative of the susceptance of a straight tube [ə] with the dimensions as used above. It can be seen that the susceptance of the air sac consists of monotonically increasing curves with vertical asymptotes at the points of zero impedance. The susceptance at higher frequencies is negative, and approximately equal for all types of air sacs. The negative susceptance of the tube consists of monotonically decreasing lines, separated by vertical asymptotes where the impedance is zero. A qualitatively similar curve is obtained for differently shaped tracts, but with differently spaced asymptotes. A minimum for admittance (and therefore a maximum for impedance) for the parallel circuit is achieved where the tract curves and the air sac curves intersect. For reference, without an air sac, resonances occur where the oral tract curve intersects the zero susceptance line (dotted horizontal line in Fig. 10). From Fig. 10 it becomes clear why the air sac changes the oral tract's response the way it does. The resonances at low frequencies occur near the intersection of the oral tract curve and the vertical asymptotes of the air sac susceptance. As the intersections' frequency is mostly determined by the position of the vertical asymptotes, it is therefore expected to be relatively independent of the oral tract shape. As for the other intersections, they are expected to occur at higher frequencies than those of the oral tract without an air sac. This is because the air sac susceptance is negative, and the slope of the oral tract susceptance curve is also negative. Because the air sac susceptance curve approaches zero for higher frequencies, it is also expected that higher-frequency resonances are shifted less than lower ones, therefore pushing them closer together.

The above analysis is only valid for low frequencies and for situations where the dimensions of the oral tract and the air sac are not such that the interactions occur at qualitatively different places. However, for all dimensions that are realistic for primates, the analysis appears to remain valid. As for higher frequencies, it can be observed in Fig. 8 that extra peaks at higher frequencies occur in the combined spectrum. These peaks correspond to resonances of the air sac, and are therefore insensitive to vocal tract shape.

III. VALIDATING THE MODEL

The simplifications made in the lumped element model appear to be valid in a mathematical sense, as it gives the same result as the finite element model. However, two things remain to be investigated: is the model physically plausible and is it able to explain properties of primate vocalizations? For this reason two short experimental studies were conducted.

For testing the accuracy of the physics, artificial models made out of perspex tubes were used. This was done because the dimensions and physical properties of these models could be measured exactly, and therefore the only uncertainty in these experiments was whether the modeled physics correspond to reality, and therefore a good quantitative correspondence is expected.

For testing the applicability of the theory to primate vocalizations, a first attempt was made to model two primates' vocal tracts and air sacs based on anatomical data, and to compare the calculated resonance frequencies with those measured in actual animal calls. The two species selected were the red howler monkey (*Alouatta seniculus*) and the siamang (*Symphalangus syndactylus*). Both have prominent air sacs as well as calls with prominent low-frequency components. The red howler monkey has a complex air sac

system with both paired lateral ventricular and subhyoidal air sacs, and most likely has the largest hollow hyoid bone of all primates. The subhyoidal air sac is situated almost completely in the hyoid bulla, while the lateral air sacs are situated in the thyrocuneiform space. The lateral ventricular and subhyoidal air sacs appear each to be connected to the larynx separately, and therefore form three separate air sacs. The subhyoid airsac has bony walls, while the lateral ventricular air sacs has partly cartilaginous walls. The whole system is described in considerable detail by Schön (1971), who proposes that the lateral air sacs do not have an acoustic function, but serve to keep the channel into the subhyoid air sac open. The siamang has a single air sac, comprised of the fusion of the paired lateral ventricular air sacs. The diagrams in Fig. 1 are simplified representations of the red howler (subhyoidal air sac only) and siamang anatomy. As there was considerable uncertainty in the anatomical data (only data from the literature was used) only a qualitative correspondence can be expected.

A. Materials and Methods

1. The perspex models

Physical model of vocal tracts with air sacs were constructed from acrylic glass (perspex) tubes with a wall thickness of 2 mm. Two sets of three models were constructed, one set with and one without air sacs for two-tube models corresponding roughly to the vowels [a], [ə] and [y] and having dimensions given in tables III and IV. The density of perspex is 1.2 kg·m^{-3}, Young's modulus is approximately 3 GPa and Poisson's ratio 0.35. The value for k_2 was therefore estimated to be 2.3 GPa. The models are illustrated in Fig. 11 A.

The resonance frequencies of the perspex models were measured by exciting them with a single pulse (generated by tapping the closed end). The sound was recorded with an Altec Lansing AHS30 electret microphone at 5 cm from the model's mouth and sampled at 44.1 KHz using a Soundmax Integrated Digital HD Audio PC card. A power spectrum was calculated with the PRAAT software package version 5.0.16 (Boersma & Weenink, 2008) using a 0.05s Hann window. The position of the resonance peaks in the spectrum was calculated as the average of measurements on three different pulses.

The air sac was modeled with the complete lumped element model and with the Finite Element model (using a mesh of 4708 triangular elements) described in section II.A.8, while the vocal tracts were modeled with Flanagan's transmission line model (Flanagan, 1965). Radiation at the lips was also modeled after Flanagan (1965). The radiated power was calculated for frequencies between 1 and 4000 Hz, and the peaks were recorded. For these calculations, the speed of sound at room temperature ($345 m \cdot s^{-1}$) was used rather than the speed of sound at body temperature as given in Table I.

2. The Animal models

The model of the red howler monkey was constructed using anatomical data (Kelemen & Sade, 1960; Schön, 1971; Schön Ybarra, 1986) and an estimated vocal tract length of 11 cm. The model consisted of a 4-tube vocal tract with an attached simple air sac. The air sac does not model the complex anatomy of the howler monkey air sac system, but consists of a hard walled resonator of 5 cm diameter, 40 cm^3 volume and 40 cm^2 (radiating) surface area, connected to the glottis with a 3 cm tube of 1 cm diameter. The howler monkey's lateral ventricular sacs were not modeled in this first approximation, as Schön Ybarra has proposed that the subhyoidal air sac is the main resonator (Schön,

1971, in the discussion). The oral tract consists of a 3 cm long tube of 3 cm diameter modeling the chamber above the glottis, a 1.5 cm long tube of 2 cm diameter modeling the constriction near the laryngeal introitus (number 7 in Schön Ybarra 1986, Fig. 5) a 5 cm long tube of 4 cm diameter modeling the oral cavity and a 1.5 cm long tube of 3 cm diameter modeling the rounded lips. This model is based on rough estimates, but it is meant to show a qualitative, rather than a complete quantitative resemblance to the howler monkey's bark. The model is illustrated in figure 11 B.

Schön Ybarra (1986) presents an analysis of calls from the red howler monkey. In that study, different types of calls (called barks and roars) occur that have resonances at around 350–400, 900–1100 and 1800–2200 Hz, although not all resonances appear in all calls. Some calls lack either or both peaks around 1000 Whitehead (1995) presents resonance patterns for roars that have been derived from a "multi-animal chorus" (multiple animals vocalizing during the same time period) with comparable values as the ones found by Schön Ybarra, and Whitehead also mentions that he does not consistently find the third peak around 2000 Hz. An example of a howler monkey's spectrum and spectrogram are illustrated in Fig. 12 A (based on Macaulay library recording #25513, recorded by Thomas H. Davis in Suriname on 23 February 1981, digitized at 44.1 KHz; the spectrum was calculated from a 2205-sample Hann window). This figure shows the same qualitative picture as found by Schön Ybarra and Whitehead, except that it appears that there are *two* resonances in the range of 700–1000 Hz. In Whitehead's spectrogram (Whitehead, 1995, figure 7), the peak appears wide, but one cannot clearly distinguish two peaks. However, an interpretation in terms of two resonances seems possible.

A second model was constructed to model the air sac of the siamang (*Symphalangus syndactylus*). In this case a larger air sac, of the dimensions given in Table I was used. Given the near absence of higher harmonics in the boom call, the oral part of the vocal tract was assumed to be closed (no radiation).However, a short laryngeal-pharyngeal section of 3 cm length and 2 cm diameter was assumed (lacking precise anatomical data about the siamang, these dimensions were based on the *Hylobates lar* larynx shown in Nishimura, 2003). Assuming a closed vocal tract also means that the air sac inflates or deflates during vocalization (depending on whether it occurs on exhalation or inhalation) and that its acoustic properties change. Therefore the spectrum presented here should be seen as a single time slice of the vocalization. The model is illustrated in figure 11 C. Siamang booms (the low frequency parts of the call, produced with the air sac inflated, Haimoff, 1981) have a clear peak around 250–300 Hz (Haimoff, 1981; Riede et al., 2008, figure 10B), illustrated in Fig. 12 B (based on Macaulay library recording #126331 recorded by Andrea L. Priori at the Washington Zoo on 22 February 1973, digitized at 44.1 KHz; the spectrum was calculated from a 2205-sample Hann window). In this recording, the frequency of the boom increases from about 245 Hz at the beginning to 260 Hz at the end, indicating that the air sac deflates during vocalization. The vocal folds probably vibrate close to the frequency of this first resonance, as second and third harmonics are visible in the spectrum and spectrogram. As the third harmonic is more prominent than the second, there must be a supraglottal resonance at around 800–900 Hz. The vocal tract was modeled with Flanagan's transmission line model, and the air sac was modeled with the complete lumped element model. The radiated power was calculated for frequencies between 1 and 4000 Hz, and the peaks were recorded.

B. Results

1. The perspex models

The measurements and the values predicted by the computer model are given in Table V and illustrated in Fig. 13. It can be observed that the differences between calculated and measured resonances are small and that the lumped element model gives qualitatively similar, but slightly less accurate predictions than the finite element model. The difference between the finite element model and the measured values are on average 2.9% with a maximum of 5.3% while for the lumped element model, the average is 6.3% and the maximum 14.1%. Variation between different measurements of the recorded signal was approximately 2%.

The results conform to the theory: a new resonance at approximately the resonance frequency of the air sac appears and the frequency of this resonance is insensitive to the shape of the vocal tract. Simultaneously, the first two resonance frequencies of the original vocal tract are shifted up and closer together.

2. The animal models

The calculated spectrum of the howler monkey model with air sac is shown in Fig. 14. Resonances occur at 369, 921 and 1827 Hz. For reference, the spectrum of the Howler monkey vocal tract without an air sac is also given, and its resonances occur at 798 and 1761 Hz. It can be observed that there are no prominent higher resonances below 3000 Hz, just as is the case in Schön Ybarra's sonogram (Schön Ybarra 1986, figure 1). There is also a reasonable qualitative similarity between the curve in Fig. 14 and the spectrogram shown in Whitehead (1995) and in Fig. 12 A. With respect to the data shown

in Fig. 12 A, it is important to note that in the real howler monkey call there are two peaks in the range of 700–1000 Hz instead of one.

The calculated radiated power of the siamang model is given as the solid line in Fig. 14. Two prominent peaks can be observed at 228 and 875 Hz, which correspond qualitatively to the siamang's boom, although in the siamang (Fig. 12 B) the resonance near 800 Hz is much less prominent. The first resonance is due to the air sac, the second to the laryngopharyngeal cavity. The small peak around 500 Hz (best visible in the spectrogram) corresponds to the second harmonic of vocal fold vibration (which if tuned at the resonance of the air sac would be around 228 Hz). As the model does not take source characteristics into account, no corresponding peak is expected in the modeled spectrum.

It should be noted that all acoustic energy in the siamang model is radiated from the air sac. For the low frequencies involved, the air sac is a much better radiator than the open mouth, as it has a larger surface area. Its function as a radiator might explain the absence of hair on the skin covering the siamang's air sac.

C. Discussion

In the best controlled case, that of the perspex models, the correspondence between the resonance frequencies calculated by the finite element model and those observed is within 5.3%, which is comparable with the uncertainties with which the resonances and the dimensions of the models could be measured. This confirms that the underlying physical analysis of air sac behavior is sound. The results also confirm the qualitative analysis of the lumped element model: an extra peak appears at a frequency that is close

to the resonance frequency of the air sac itself, while the original resonances of the vocal tract to which the air sac is attached, are shifted up and closer together.

As for the animal models, the uncertainties here are much bigger, as only limited anatomical data was available. The correspondence between the howler monkey's call and the calculated spectrum is incomplete. The three predicted resonances occur at approximately the right frequencies, but it appears as if the real howler monkey spectrum has an extra resonance in the 700–1000 Hz range. It is unlikely that this is due to the shape of the vocal tract that was modeled. A different shape of the vocal tract could result in a different pattern of resonances, but the shape of the vocal tract investigated here already results in resonances that have low frequencies for a vocal tract with a length of 11 cm. Shapes with as many low resonances as observed would require a vocal tract area function with unrealistically sharp transitions. It is more likely that the extra resonance is caused by the lateral air sacs of the howler monkey. These structures are not taken into account in the present model, but might produce extra resonances in the correct frequency range (given a very rough size estimate derived from Schön (1971) Fig. 6). It therefore appears that Schön Ybarra's (Schön, 1971, in discussion) remark that the "...bulla hyoidea can act as the main resonator..." is mostly correct, but that the lateral sacs may cause an extra resonance. This should be investigated with a more complete model, but that requires accurate anatomical data about the howler monkey's lateral air sacs. Using accurate anatomical data is critical, as the match between the animal's resonances spectrum and the model's resonances can always be improved by tuning of parameters, such as the dimensions of the vocal tract and the air sac. Such parameter tuning should be avoided, however, as one can easily fool oneself by constructing a model that shows

similar resonances as the animal model, but for totally different acoustic reasons. A more anatomically correct model should also be able to explain how the variations between the different howler calls can be made, potentially by attaching and detaching the different air sacs, and by opening and closing the vocal tract.

The qualitative correspondence between the predicted and observed siamang spectra is good. The only important difference is that in the real siamang spectrum, the second resonance is much less prominent. This could be due to coupling between the source of vocal energy and the air sac, which causes the source to contain much less energy at the second resonance, but this should be investigated further with more accurate models.

IV. CONCLUSION

The acoustic effect of air sacs with (to first order) constant volume has been investigated in this paper. It was shown that only the inertance of the air in the neck of the air sac, the volume of air in the sac itself and the mass and stiffness of the wall of the air sac have significant influence on the resonance frequencies. Whether mass or stiffness dominates depends on the air sac's wall tissue. For soft tissue, mass dominates. For bony tissue, stiffness dominates. For cartilaginous tissue both mass and stiffness need to be taken into account. It was also shown that for large air sacs with flexible walls, radiation of acoustic power through the walls makes a significant contribution to the total sound produced at low frequencies.

As for the spectrum of the sound produced by a vocal tract with an attached air sac, it has been shown that the presence of a large air sac results in extra resonances at the approximate resonance frequencies of the isolated air sac. Attaching an air sac to a given oral tract also results in a shifting up of the original resonances (at least for resonances

lower than the second resonance frequency of the air sac). As lower resonances are shifted up more than higher resonances, this process results in resonances moving closer together. These findings correspond to the experimental and modeling results of (Riede et al., 2008).

The theoretical model was able to predict the resonances of a perspex model accurately. Moreover, models based on the anatomy of a red howler monkey and of a siamang resulted in resonance patterns that were qualitatively (and to some degree quantitatively) similar to those of certain vocalizations produced by these species. The major resonances of the siamang "boom" could be modeled correctly, whereas three of the four major resonances of the red howler monkey could be modeled correctly by only modeling its bony-walled subhyoidal air sac. It was proposed that for a complete howler monkey model, the lateral air sacs also need to be taken into account. The model developed in this paper can be used for the case of multiple air sacs, by adding for each air sac a lumped element circuit as an extra parallel impedance. The results support the acoustical model developed in this paper, as well as the hypothesis that air sacs have an acoustic function in howler monkeys and siamangs.

Only the properties of the air sac as an acoustic filter have been investigated, while it was assumed that the source of vocal energy is a perfect volume velocity source. This simplification was made in order to enable theoretical analysis. However, interaction between the vocal folds and the vocal tract is probably very important in primate calls. Riede et al. (2008) have found such interactions between the air sac and the vocal folds in an artificial model. Even in the absence of source-filter interaction, it would make a large difference in radiated power at what exact frequency the vocal folds vibrate. Tuning the

fundamental frequency to fall near an anti-resonance of the air sac might, on the contrary, help to suppress its first harmonic. This would concentrate acoustic output at higher frequencies, something which has been proposed for the high-frequency calls of siamangs by Haimoff (1983).

Finally, it is tempting to speculate about the influence that air sacs would have on speech, and their possible relevance for the study of the evolution of speech. The presence of an air sac results in a resonance at low frequencies, whose frequency is insensitive to the shape of the oral tract. In addition, higher resonances get shifted up and closer together. As distinctions between speech sounds are perceived primarily on the basis of the differences between their resonance patterns, this means that a smaller range of distinctive speech sounds could be produced when an air sac is present. This could be investigated by determining the maximal vowel space (Boë *et al.*, 1989) with and without an air sac, in a way similar to what has been done by Boë et al. (2002) and de Boer (2008b) and.

In this context, it would be interesting to investigate the perceptual effect of connecting an air sac to a vocal tract. This could be done with psychoperceptual experiments, in which the ability to distinguish between different vocal tract shapes with and without air sacs is investigated. Such investigations, however, fall outside the scope of this paper. A first impression of the perspex models with and without air sacs is that of a clear decrease in perceived pitch when an air sac is attached to a vocal tract, while the different vocal tracts with an air sac sound more similar than those without air sacs.

A potential effect of the presence of the lower resonance is that it may help to exaggerate size. Without air sacs, the resonance patterns produced by the models with air sacs could

only be produced by much longer vocal tracts. This can be an evolutionary advantage (Fitch, 2000; Fitch & Hauser, 1995, 2002). However, when being able to produce a wide range of distinctive resonance patterns becomes evolutionarily important, such as must have happened in the evolution of speech, air sacs become a disadvantage. Even if it is possible that air sacs can be disconnected from the vocal tract (which is likely) they can nevertheless be a source of infection, called airsacculitis (Lawson et al., 2006, and references therein). It therefore makes sense that they would disappear if they are no longer useful in vocalization.

Although the model has been developed for understanding the function of air sacs in primates and ancestral hominins, there is no reason that the model could not be applied to other animals with air sacs, as long as their structure consists of a neck (a tube) connected to a cavity. This appears to be the case in many other animals. Frey et al. (2007) study airsacs in reindeer and give an overview of many mammalian species, while for example Reidenberg and Laitman (2008) give an overview of cetacean air sacs and Weissengruber et al. (2001) give an overview of epipharyngeal pouches (equivalent to air sacs) in bears. The model should be applicable to understanding these animals' vocalizations.

In any case a better understanding of the acoustic function of air sacs, through analysis of simplified models, through extension of the theory to include source-filter interactions and through experiments such as presented here and in Riede et al. (2008) will help to advance the understanding of the evolutionary and behavioral function of these fascinating anatomical structures.

Acknowledgements

This work is part of the NWO vidi project "Modeling the evolution of speech" grant number 016.074.324. The author thanks Didier Demolin for discussion on the function and anatomy of air sacs, Rob van Son for comments on the manuscript and Wendy van Bohemen for letting me use the anatomical collection of the Amsterdam zoological museum.

References

Alemseged, Z., Spoor, F., Kimbel, W. H., Bobe, R., Geraads, D., Reed, D., et al. (**2006**). A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature, 443*(7109), 296–301.

Alemseged, Z., Spoor, F., Kimbel, W. H., Bobe, R., Geraads, D., Reed, D., & Wynn, J. G. (**2006**). A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature, 443*(7109), 296–301.

Arensburg, B., Tillier, A. M., Vandermeersch, B., Duday, H., Schepartz, L. A., & Rak, Y. (**1989**). A middle Palaeolithic human hyoid bone. *Nature, 338*(6218), 758–760.

Avril, C. (**1963**). Kehlkopf und Kehlsack des Schimpansen, *Pan troglodythes* (Blumenbach 1799). (mamalia, primates, Pongidae). (Larynx and air sac of the chimpanzee, *Pan troglodytes* (Blumenbach 1799). (Mamalia, Primates, Pongidae)) *Gegenbaurs morphologisches Jahrbuch, 105*, 75–129.

Blackstock, D. T. (**2000**). *Fundamentals of physical acoustics*. New York: John Wiley & sons, inc.

Boë, L.-J., Heim, J.-L., Honda, K., & Maeda, S. (**2002**). The potential Neandertal vowel space was as large as that of modern humans. *Journal of Phonetics, 30*(3), 465–484.

Boë, L.-J., Perrier, P., Guerin, B., & Schwartz, J.-L. (**1989**). *Maximal vowel space.* Paper presented at the Eurospeech, Paris, France.

Boersma, P., & Weenink, D. (**2008**). Praat: Doing phonetics by computer (Version 5.0.16). Amsterdam.

Brandes, R. (**1932**). Über den Kehlkopf des Orang-utan in verschiedenen Altersstadien mit Berücksichtigung des Kehlsackfrage. (About the larynx of the orangutan in different stages of development, with discussion of the matter of air sacs) *Gegenbaurs morphologisches Jahrbuch, 69*(1–61).

Camper, P. (**1779**). Account of the ogans of speech of the orang outang. By Peter Camper, M. D. Late professor of anatomy, &c. In the university of Groningen, and f. R. S. In a letter to sir John Pringle, f. R. S. *Philosophical Transactions of the Royal Society of London, 69*, 139–159.

Dang, J., & Honda, K. (**1997**). Acoustic characteristics of the piriform fossa in models and humans. *Journal of the Acoustical Society of America, 101*(1), 456–465.

Dang, J., Honda, K., & Suzuki, H. (**1994**). Morphological and acoustical analysis of the nasal and paranasal cavities. *Journal of the Acoustical Society of America, 96*(4), 2088–2100.

de Boer, B. (**2008a**). The acoustic role of supralaryngeal air sacs. *Journal of the Acoustical Society of America, 123*(5 Pt 2), 3732–3733.

de Boer, B. (**2008b**). The joy of sacs. In A. D. M. Smith, K. Smith & r. Ferrer i Cancho (Eds.), *The evolution of language* (pp. 415–416). New Jersey: World Scientific.

Dular, P., Geuzaine, C., Henrotte, F., & Legros, W. (**1998**). A general environment for the treatment of discrete problems and its application to the finite element method. *IEEE Transactions on Magnetics, 34*(5), 3395–3398.

Fant, G. (**1960**). *Acoustic theory of speech production*. 'sGravenhage: Mouton.

Fant, G. (**1972**). Vocal tract wall effects, losses, and resonance bandwidths. *Speech Transmission Laboratory Quarterly Progress and Status Report, 13*(2–3), 28–52.

Fitch, W. T. (**2000**). The evolution of speech: A comparative review. *Trends in cognitive sciences, 4*(7), 258–267.

Fitch, W. T., & Hauser, M. D. (**1995**). Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on "honest" advertisement. *American Journal of Primatology, 37*(3), 191–219.

Fitch, W. T., & Hauser, M. D. (**2002**). Unpacking "honesty": Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, R. R. Fay & A. N. Popper (Eds.), *Acoustic communication* (pp. 65–137). New York: Springer.

Flanagan, J. L. (**1965**). *Speech analysis, synthesis and perception*. Berlin: Springer.

Fletcher, N. H., Riede, T., Beckers, G. J. L., & Suthers, R. A. (**2004**). Vocal tract filtering and the ‘‘coo’’ of doves. *Journal of the Acoustical Society of America, 116*(6), 3750–3756.

Frey, R., Gebler, A., Fritsch, G., Nygrén, K., & Weissengruber, G. E. (**2007**). Nordic rattle: The hoarse vocalization and the inflatable laryngeal air sac of reindeer (*Rangifer tarandus*). *Journal of Anatomy, 210*(2), 131–159.

Gautier, J.-P. (**1971**). Etude morphologique et fonctionelle des annexes extra-laryngées des Cercopithecinae; liaison avec les cris d'espacement. (Morphological and functional study of the extralaryngeal sacs of the Cercopithecines; relation to separation calls) *Biologica Gabonica, 7*, 229–267.

Givoli, D. (**1991**). Non-reflecting boundary conditions. *Journal of Computational Physics, 94*, 1–29.

Haimoff, E. F. (**1981**). Video analysis of siamang (*Hylobates syndactylus*) songs. *Behaviour, 76*(1), 128–151.

Haimoff, E. F. (**1983**). Occurence of anti-resonance in the song of siamang (*Hylobates syndactylus*). *American Journal of Primatology, 5*(3), 249–256.

Hewitt, G. P., MacLarnon, A., & Jones, K. E. (**2002**). The functions of laryngeal air sacs in primates: A new hypothesis. *Folia Primatologica, 73*, 70–94.

Hilloowala, R. A., & Lass, N. J. (**1978**). Spectrographic analysis of laryngeal air sac resonance in rhesus monkey. *American Journal of Physical Anthropology, 49*(1), 129–132.

Hunter, E. J., & Titze, I. R. (**2007**). Refinements in modeling the passive properties of laryngeal soft tissue. *Journal of Applied Physiology, 103*(1), 206–219.

Jin, H., & Lewis, J. L. (**2003**). *Determination of Poisson's ratio of articular cartilage in indentation test using different sized indenters.* Paper presented at the Summer Bioengineering Conference, Key Biscayne (FL).

Kelemen, G., & Sade, J. (**1960**). The vocal organ of the howling monkey (*Alouatta palliata*). *Journal of Morphology, 107*(2), 123–140.

Kleinschmidt, A. (**1938**). Die Schlund- und Kehlorgane des Gorillas "Bobby" unter besonderer Berücksichtigung der gleichen Organe von Mensch und Orang. Ein beitrag zur vergleichenden Anatomie des Kehlkopfes. (The pharyngeal and laryngeal organs of the gorilla “Bobby” with special discussion of the equivalent organs of humans and orangs. A contribution to the comparative anatomy of the larynx.) *Gegenbaurs morphologisches Jahrbuch, 81*, 78–157.

Lambert, R. K., Baile, E. M., Moreno, R., Bert, J., & Paré, P. D. (**1991**). A method for estimating the Young’s modulus of complete tracheal cartilage rings. *Journal of Applied Physiology, 70*(3), 1152–1159.

Lawson, B., Garriga, R., & Galdikas, B. M. F. (**2006**). Airsacculitis in fourteen juvenile southern bornean orangutans (*Pongo pygmaeus wurmbii*). *Journal of Medical Primatology, 35*(3), 149–154.

Maeda, S. (**1982**). A digital simulation method of the vocal-tract system. *Speech Communication, 1*(3-4), 199–229.

Martin, R. B., Burr, D. B., & Sharkey, N. A. (**1998**). *Skeletal tissue mechanics*. New York: Springer-Verlag.

Martínez, I., Arsuaga, J.-L., Quam, R., Carretero, J.-M., Gracia, A., & Rodríguez, L. (**2008**). Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *Journal of Human Evolution, 54*, 118–124.

Negus, V. E. (**1949**). *The comparative anatomy and physiology of the larynx*. London: William Heinemann Medical Books Ltd.

Nishimura, T. (**2003**). Comparative morphology of the hyo-laryngeal complex in anthropoids: Two steps in the evolution of the descent of the larynx. *Primates, 44*(1), 41–49.

Nishimura, T. (**2008**). Origin of human speech an primate vocalizations: Paleoanthropology and bioacoustics. *Anthropological Science (Japanese Series), 116*(1), 1–14.

Nishimura, T., Mikami, A., Suzuki, J., & Matsuzawa, T. (**2007**). Development of the laryngeal air sac in chimpanzees. *International Journal of Primatology, 28*(2), 483–492.

Reidenberg, J. S., & Laitman, J. T. (**2008**). Sisters of the sinuses: Cetacean air sacs. *The Anatomical Record, 291*(11), 1389–1396.

Riede, T., Tokuda, I. T., Munger, J. B., & Thomson, S. L. (**2008**). Mammalian laryngseal air sacs add variability to the vocal tract impedance: Physical and computational modeling. *Journal of the Acoustical Society of America, 124*(1), 634–647.

Schön, M. A. (**1971**). The anatomy of the resonating mechanism in howling monkeys. *Folia Primatologica, 15*(1-2), 117–132.

Schön Ybarra, M. A. (**1986**). Loud calls of adult male red howling monkeys (*Alouatta seniculus*). *Folia Primatologica, 47*(4), 204–216.

Stevens, K. N. (**1999**). *Acoustic phonetics*. Cambridge (MA): MIT Press.

Titze, I. R. (**2002**). Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *Journal of the Acoustical Society of America, 111*(1 Pt 1), 367–376.

Titze, I. R. (**2008**). Nonlinear source–filter coupling in phonation: Theory. *Journal of the Acoustical Society of America, 123*(5), 2733–2749.

Weissengruber, G. E., Forstenpointner, G., Kübber-Heiss, A., Riedelberger, K., Schwammer, H., & Ganzberger, K. (**2001**). Occurrence and structure of epipharyngeal pouches in bears (Ursidae). *Journal of Anatomy, 198*(3), 309–314.

Whitehead, J. M. (**1995**). Vox Alouattinae: A prelimary survey of the acoustic characteristics of long-distance calls of howling monkeys. *International Journal of Primatology, 16*(1), 121–144.

TABLE I: Constants and parameters used in the lumped element model. Constants taken from (Flanagan, 1965, section 3.25).

	description	value		description	value
c	sound speed	350 $m \cdot s^{-1}$	r	sac radius	0.05 m
ρ_a	air density	1.14 $kg \cdot m^{-3}$	d	wall thickness	5×10^{-3} m
c_p	air specific heat	$1.00 \times 10^3 J \cdot kg^{-1} K^{-1}$	μ	air viscosity	18.6×10^{-6} Pa·s
λ	air heat conduction	0.023 $W \cdot m^{-1} K^{-1}$	ρ_w	wall density	10^3 $kg \cdot m^{-3}$
k_2	wall elasticity	9×10^3 soft 4.5×10^5 cartilage 15×10^9 bone	Q	wall quality factor	1 soft tissue 10 cartilage 10 bone
η	air adiabatic const.	1.4	l_e	neck length	0.03 m
A_{min}	neck area	$2.5\pi \times 10^{-5}$ m^2	S_{max}	neck circumference	$10\pi \times 10^{-3}$ m

TABLE II: Frequencies (in Hertz) of the lower peaks of tracts with air sacs, and of the first and second resonances of tracts without and with air sacs, as well as the shifts in frequency, caused by the soft-walled air sac. Peaks caused by the air sac are indicated with F_{sac}, peaks caused by the vocal tract with F.

	F_{sac} 1	F_{sac} 2	F_1	shift	F_2	shift
[a]	—	—	724	—	1334	—
$[a]_{soft}$	—	223	838	114	1424	90
$[a]_{cart}$	47	256	838	114	1424	90
$[a]_{bone}$	89	—	837	113	1424	90
[ə]	—	—	524	—	1572	—
$[ə]_{soft}$	—	225	620	96	1601	29
$[ə]_{cart}$	50	257	620	96	1601	29
$[ə]_{bone}$	95	—	620	96	1601	29
[y]	—	—	312	—	1805	—
$[y]_{soft}$	—	220	412	100	1816	11
$[y]_{cart}$	47	252	413	101	1816	11
$[y]_{bone}$	88	—	410	98	1816	11

TABLE III: Dimensions of the perspex models without air sacs, used for the measurements. Sections are numbered from the glottis.

Model	**Length 1**	**Area 1**	**Length 2**	**Area 2**
[a]	7.7 cm	2.0 cm^2	7.8 cm	10.2 cm^2
[ə]	15.8 cm	5.1cm^2	–	–
[y]	7.8 cm	10.2 cm^2	7.7 cm	2.0 cm^2

TABLE IV: Dimensions of the perspex models with air sac attachment. Tubes are numbered from the glottis. The column labelled branch indicates the distance between the glottis and the center of the tube connecting the air sac. Due to the connection between the air sac and the vocal tracts, the neck consists of two tubes with numbers 0 and 1. The air sac body is tube number 2.

Model	**Len. 0**	**Area 0**	**Branch**	**Len. 1**	**Area 1**	**Len. 2**	**Area 2**
[a]	–	–	1.2 cm	7.8 cm	2.0 cm^2	7.8 cm	10.2 cm^2
[ə]	–	–	1.4 cm	15.8 cm	5.1cm^2	–	–
[y]	–	–	1.8 cm	7.4 cm	10.2 cm^2	7.8 cm	2.0 cm^2
Air sac	1.9 cm	1.1 cm^2	–	2 cm	2.0 cm^2	8.7 cm	10.2 cm^2

TABLE V: Values for resonance frequencies of the perspex models as calculated by the lumped element model, the finite element model and as measured from the real models. For models without air sacs, only the transmission line model is used, so only one calculated value and the measured value are given. For models with air sacs, the values are given in the order: lumped element model / finite element model/ measured value.

	Fsac	**F1**	**F2**
[a] no sac		788 / – / 740	1366 / – / 1307
[a] with sac	222 / 214 / 203	888 / 866 / 891	1630 / 1478 / 1429
[ə] no sac	—	525 / – / 522	1577 / – / 1580
[ə] with sac	233 / 224 / 218	729 / 681 / 690	1579 / 1536 / 1601
[y] no sac	—	281 / – / 284	1833 / – / 1804
[y] with sac	179 / 176 / 170	567 / 511 / 499	1834 / 1815 / 1833

Figure captions:

FIG. 1 Illustration of subhyoidal air sac (left) and lateral ventricular air sac (right) and roughly representing the anatomy of the red howler monkey and siamang, respectively. Note that the howler monkey in reality also has two lateral ventricular sacs, but these are omitted to better illustrate the subhyoidal air sac. Also note that in the siamang the two lateral ventricular air sacs have merged to one sac (the septum is a vestige of the wall between the two sacs).

FIG. 2 Schematic representation of air sac, with modeled factors on the left and the lumped element electrical circuit on the right. Elements that are non-negligible are indicated in boldface. Values for elements of the circuit are given in the text.

FIG. 3 Free body diagram for wall motion. Note that although damping b and stiffness k are generated inside the wall, only the component parallel to the wall movement (due to the wall's curvature) is considered.

FIG. 4 Derivation of k_2 from generalized Hooke's law. Part A shows the equations for generalized Hooke's law, where $\sigma_{x,y,z}$ are the stresses (force per unit area) along the three axes. Part B shows its application to one-dimensional stretching, and obtains one-dimensional Hooke's law. Part C shows its application to two-dimensional stretching with constant strain in all stretching directions and obtains the relation between Young's modulus, Poisson's ratio and k_2.

FIG. 5 Impedances for frequencies between 100 and 2000 Hz for the different elements of the lumped element model. The values of C_w (dotted line) and R_w (solid line) in the 3rd graph from left are given (from bottom to top) for soft, cartilaginous and bony tissue.

FIG. 6: Absolute value of the impedances of an example air sac of the dimensions given in table I with soft, cartilaginous and bony walls. Both the horizontal and vertical scale are logarithmic. Note the peaks and the zeroes.

FIG. 7: Impedances of the finite element air sac models. Air sac dimensions are as given in table I. Note the similarity for lower frequencies to figure 6. Also note that extra pairs of peaks and zeros, as well as extra isolated peaks appear at higher frequencies in a very similar way for the three wall types.

FIG. 8: Radiated power (given a constant volume velocity source, and including power radiated from the air sac and the mouth) of three different vocal tracts. From above to below, [a], [ə] and [y] are shown. Fat grey dotted lines indicate tracts without air sacs. Solid grey lines indicate tracts with soft-walled air sacs, dashed black lines tracts with cartilaginous air sacs and solid thin black lines tracts with bony-walled air sacs. Exact dimensions are given in the text. Black lines represent tracts with air sacs, grey lines represent tracts without air sacs.

FIG. 9: Ratio of power radiated through the air sac to power radiated through the mouth for the three different wall types and for air sacs connected to the straight vocal tract.

Note that at most frequencies radiation through the mouth dominates, but that at the low frequency peaks, radiation through the air sac dominates for soft and cartilaginous walls.

FIG. 10: Nomogram to calculate the resonances of an oral tract (a straight tube) with an air sac. The narrow grey line represents the negative of the vocal tract susceptance, the broad grey line the soft tissue air sac susceptance, the dashed black line the cartilaginous air sac susceptance and the solid black line the bony air sac susceptance. The dotted line represents zero susceptance (absence of a parallel impedance). Resonances occur where the lines intersect. Tract and air sac dimensions are given in the text.

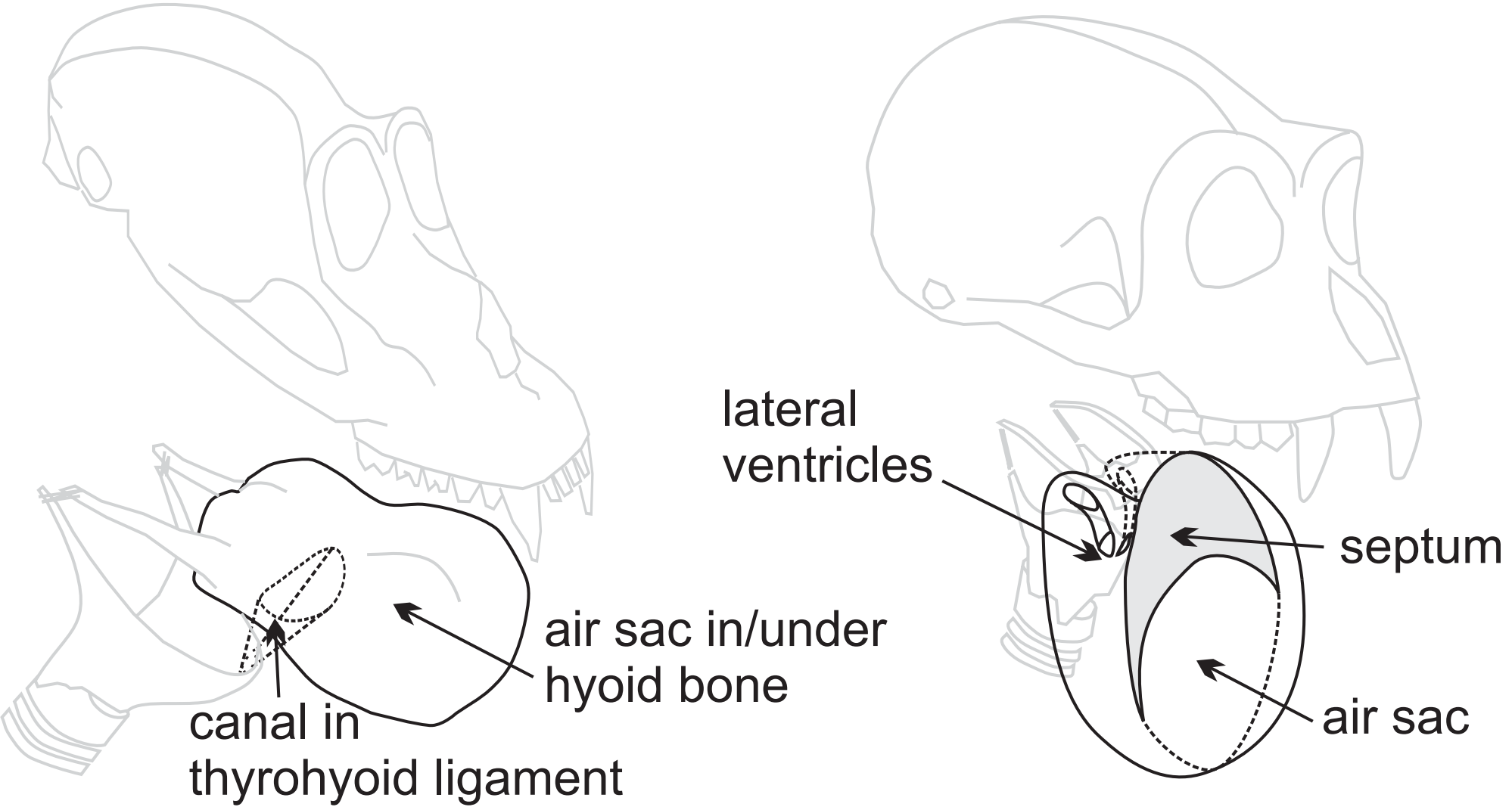
FIG. 11: Models used in the validation experiment. A: perspex models. The vertically displayed model is the air sac. The horizontally displayed models are the vocal tracts. Top row are the [a] models, middle the [ə] models and bottom the [y] models. Models without air sac are on the left, and models with air sac attachment on the right. B: red howler monkey model (the air sac is not spherical because the ratio between area and volume does not correspond to a sphere; at the frequencies involved, exact shape does not matter, only the area and the volume are relevant). C: siamang model. All models are to scale. Vocal tracts are depicted horizontally and air sacs vertically. For all models, the glottis is on the right, and the mouth is on the left. Note that the siamang model is closed at the mouth.

FIG. 12: Spectra and spectrograms of A) the red howler monkey and B) the siamang. Resonances are indicated with arrows. It is indicated in the spectrograms from which part

of the signal the spectrum was calculated. Note that in the paper, only the siamang “boom” is investigated, while in the spectrogram, also the high pitched part of the call (the “shriek”, not analyzed here) is shown (to the right). Also note that there is considerable variation in the howler monkey call (especially in the presence or absence of the resonance around 1700 Hz), while only one part of its call is modeled.

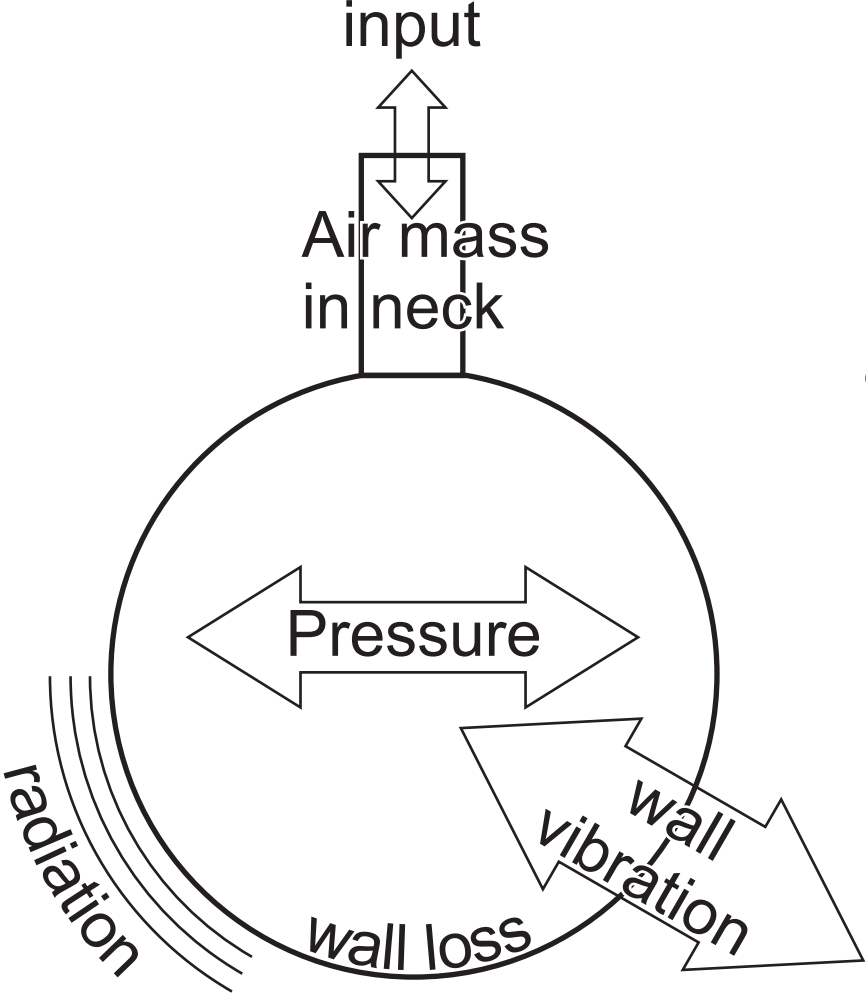
FIG. 13: Comparison of calculated (solid lines) and measured (grey lines) spectra of the perspex models without (left column) and with air sacs (right column).

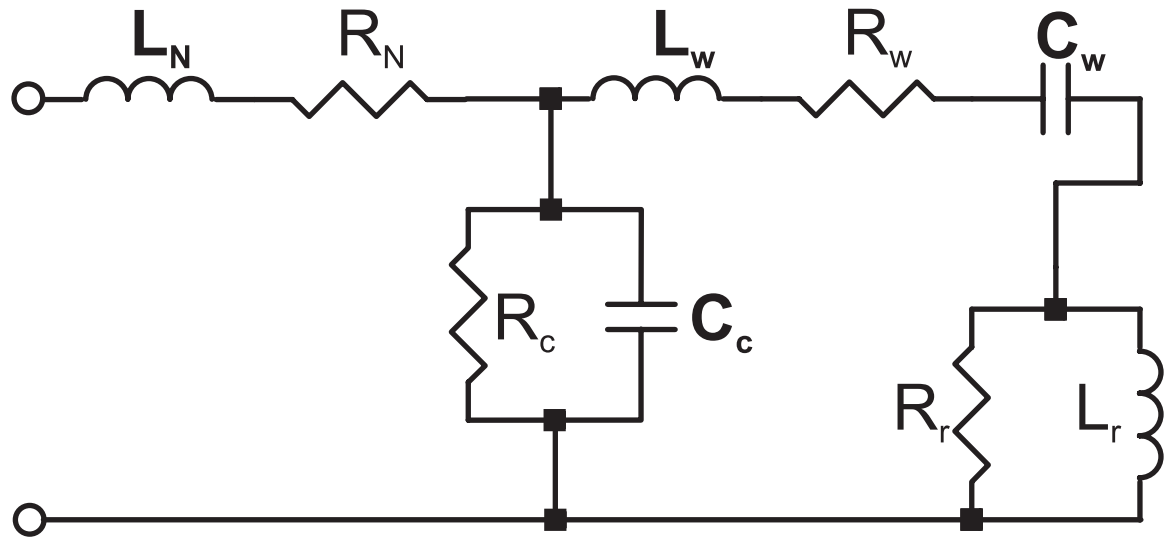
FIG. 14: Radiated power of the "howler monkey" (dashed line) and the “siamang” (solid line) models. For reference, the radiated power of a howler vocal tract without an air sac is also given (grey line). As in the siamang radiation was exclusively through the air sac, only radiated power from the vocal tract with air sac can be depicted. Details about the models can be found in the text.

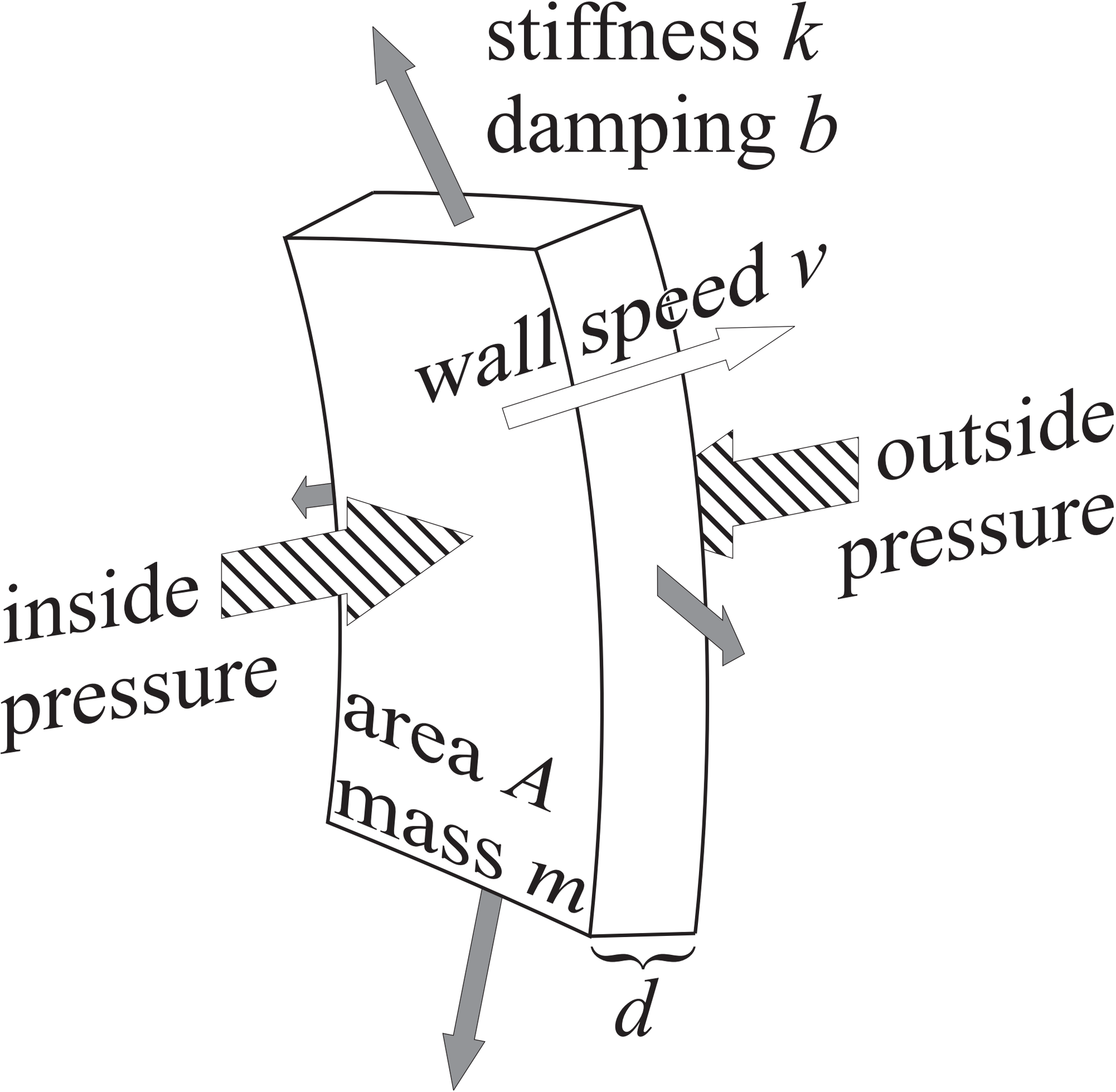


Alouatta Seniculus

Symphalangus syndactylus


input
Air mass
in neck
Pressure
radiation
wall loss
wall
vibration


L_N
R_N
L_w
R_w
C_w
R_c
C_c
R_r
L_r


stiffness k
damping b
wall speed v
outside pressure
inside pressure
area A
mass m
d

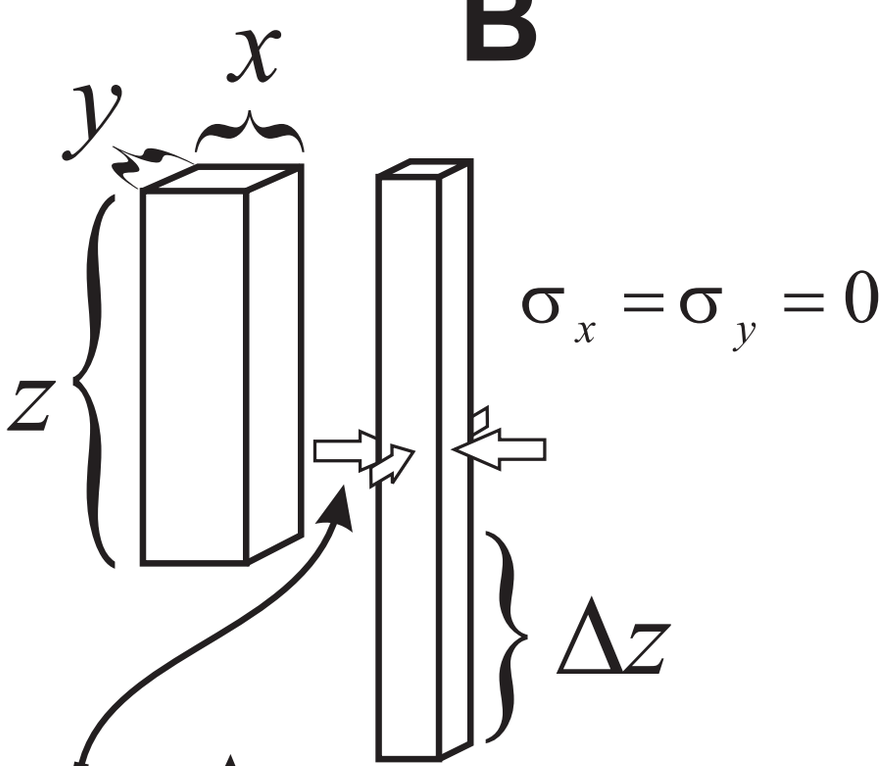
A

$$\frac{\Delta x}{x} = \frac{1}{E}\left[\sigma_x - \nu\left(\sigma_y + \sigma_z\right)\right]$$

$$\frac{\Delta y}{y} = \frac{1}{E}\left[\sigma_y - \nu\left(\sigma_x + \sigma_z\right)\right]$$

$$\frac{\Delta z}{z} = \frac{1}{E}\left[\sigma_z - \nu\left(\sigma_x + \sigma_y\right)\right]$$

B

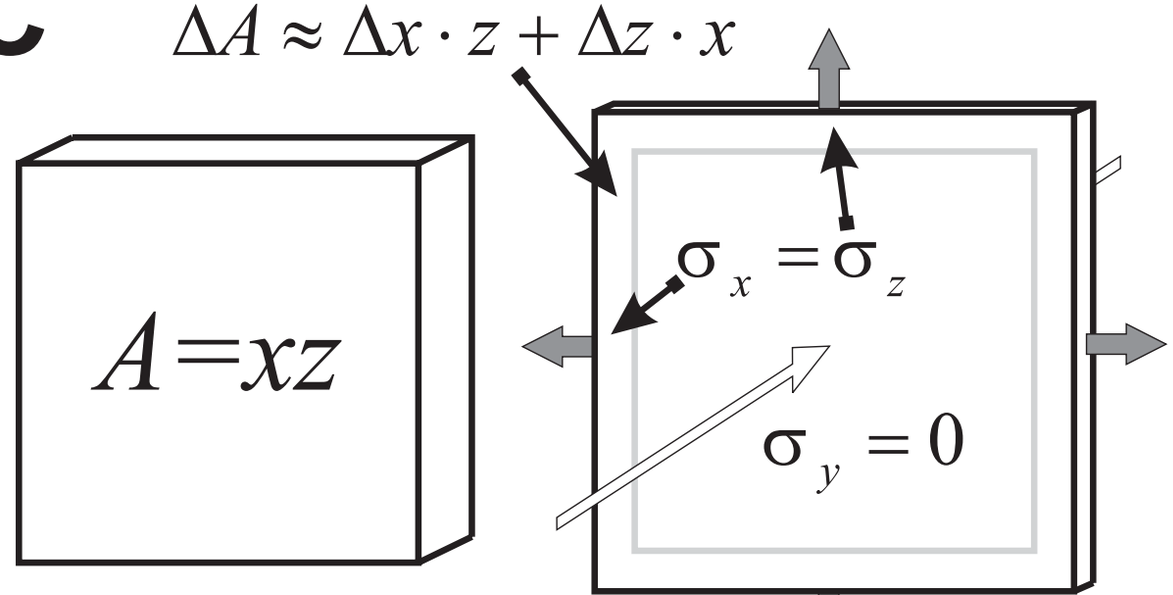


$$\frac{\Delta x}{x} = \frac{\Delta y}{y} = -\nu\,\frac{\Delta z}{z}$$

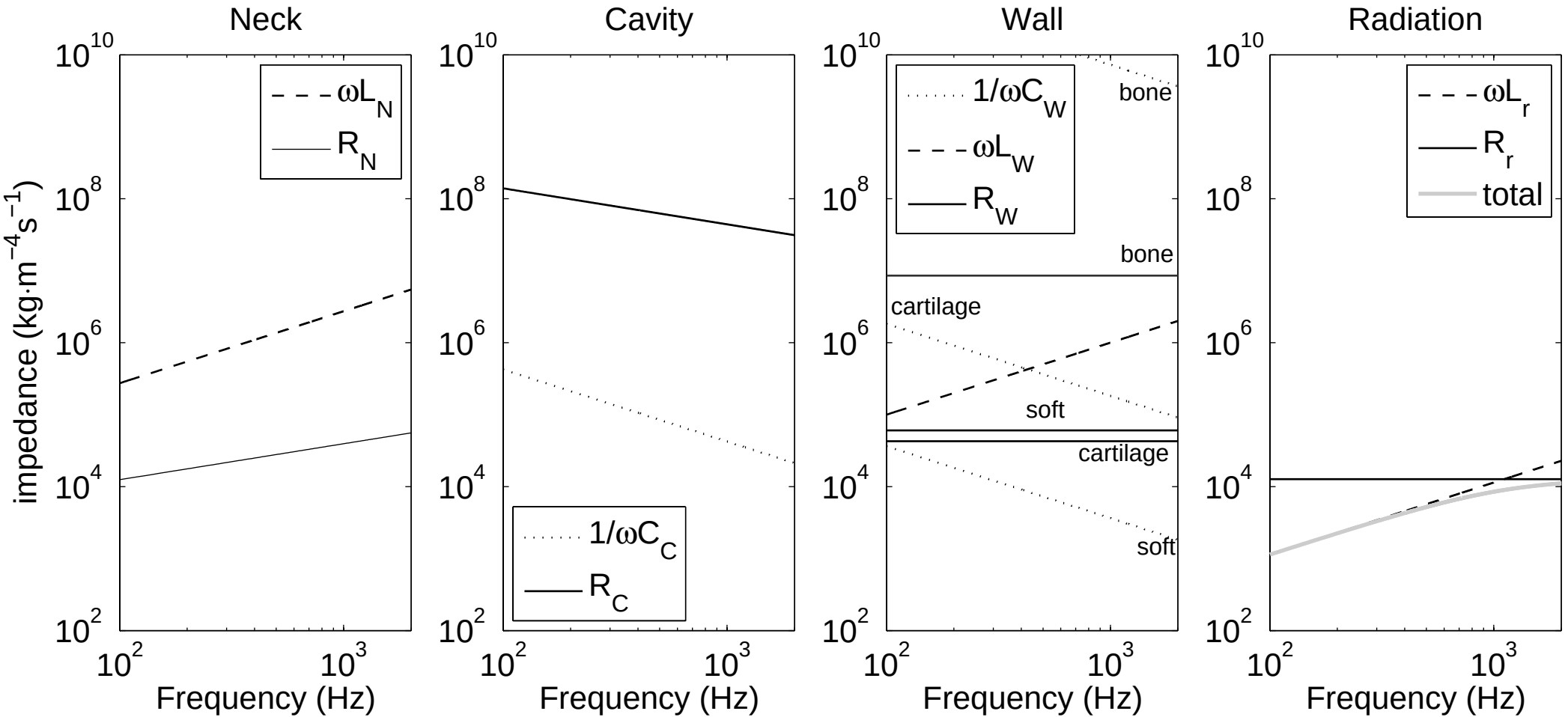
$$\sigma_z = E\,\frac{\Delta z}{z}$$

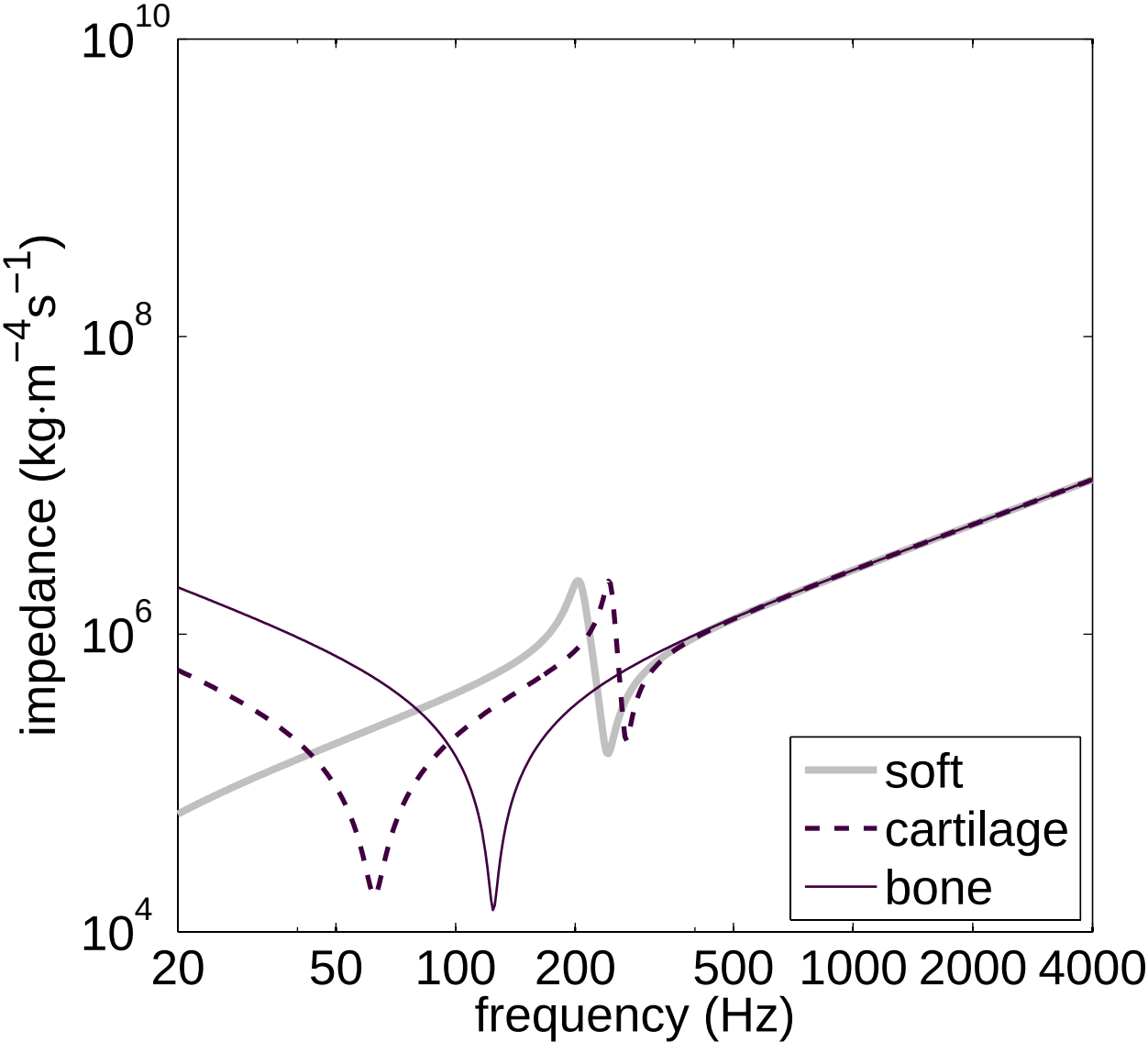
C

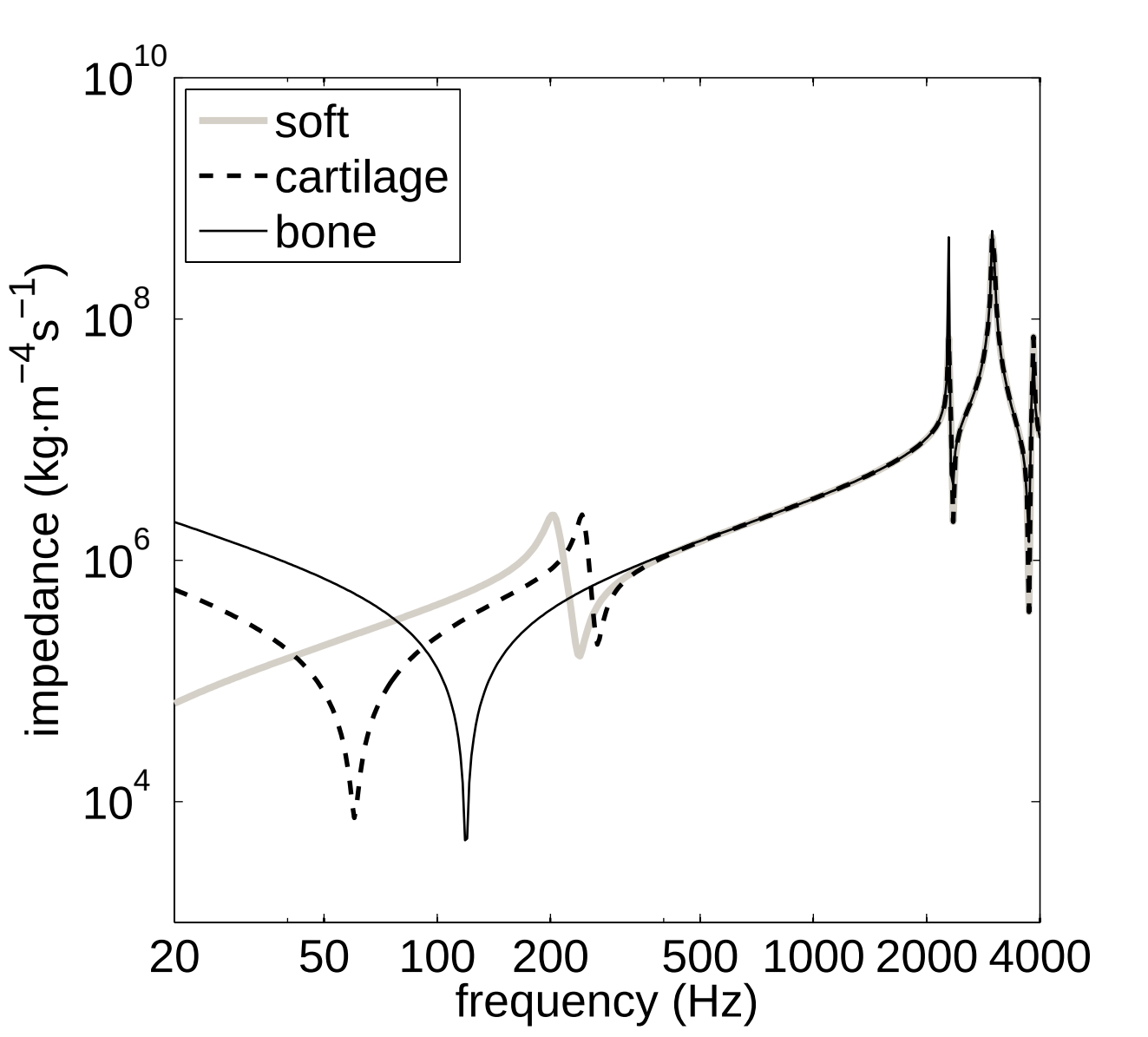
$$\Delta A \approx \Delta x \cdot z + \Delta z \cdot x$$

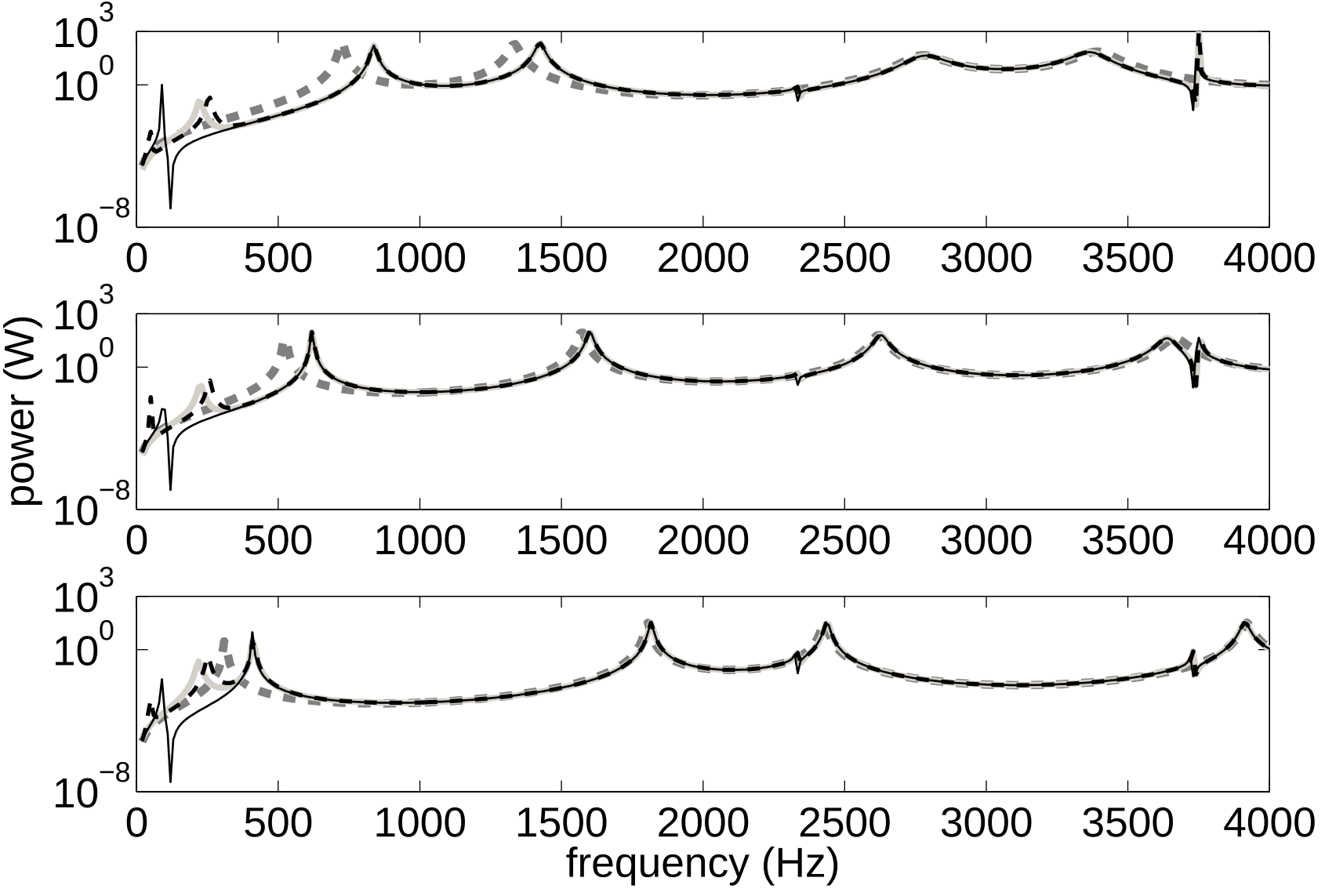


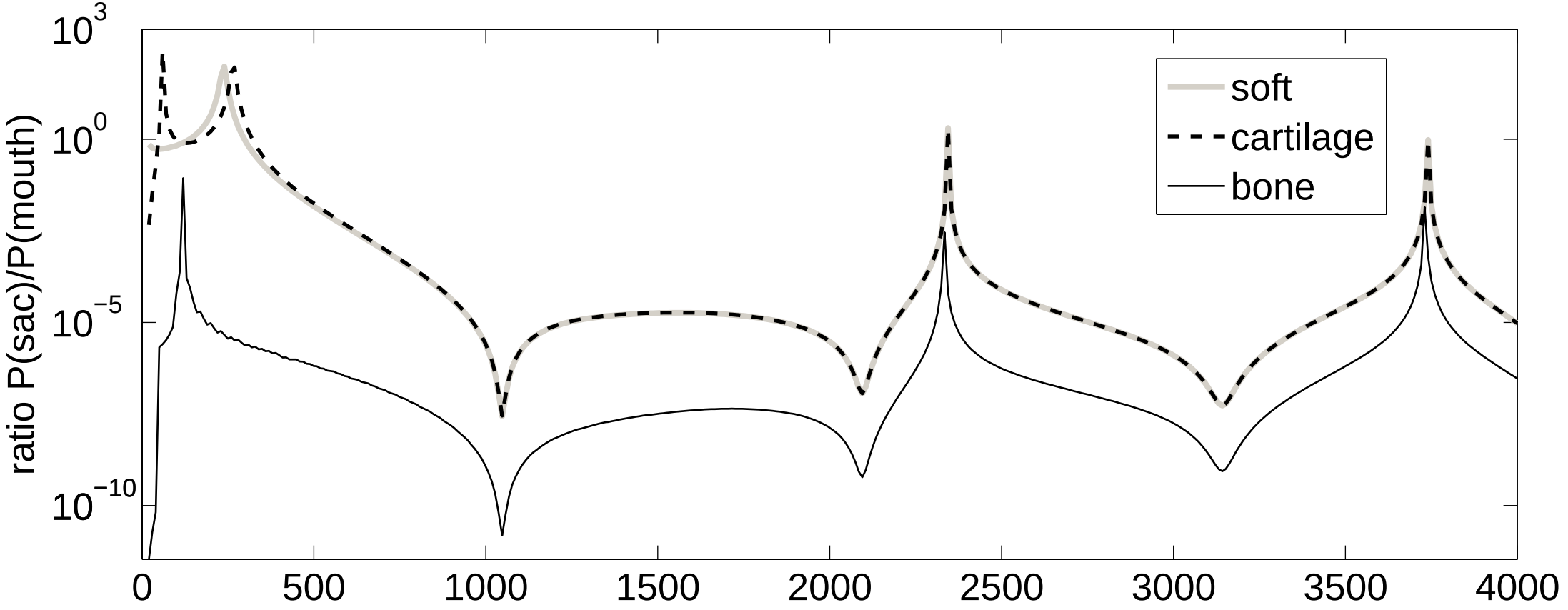
$$\sigma_x = \sigma_z = \frac{E}{2(1-\nu)}\,\frac{\Delta A}{A} = k_2\,\frac{\Delta A}{A}$$

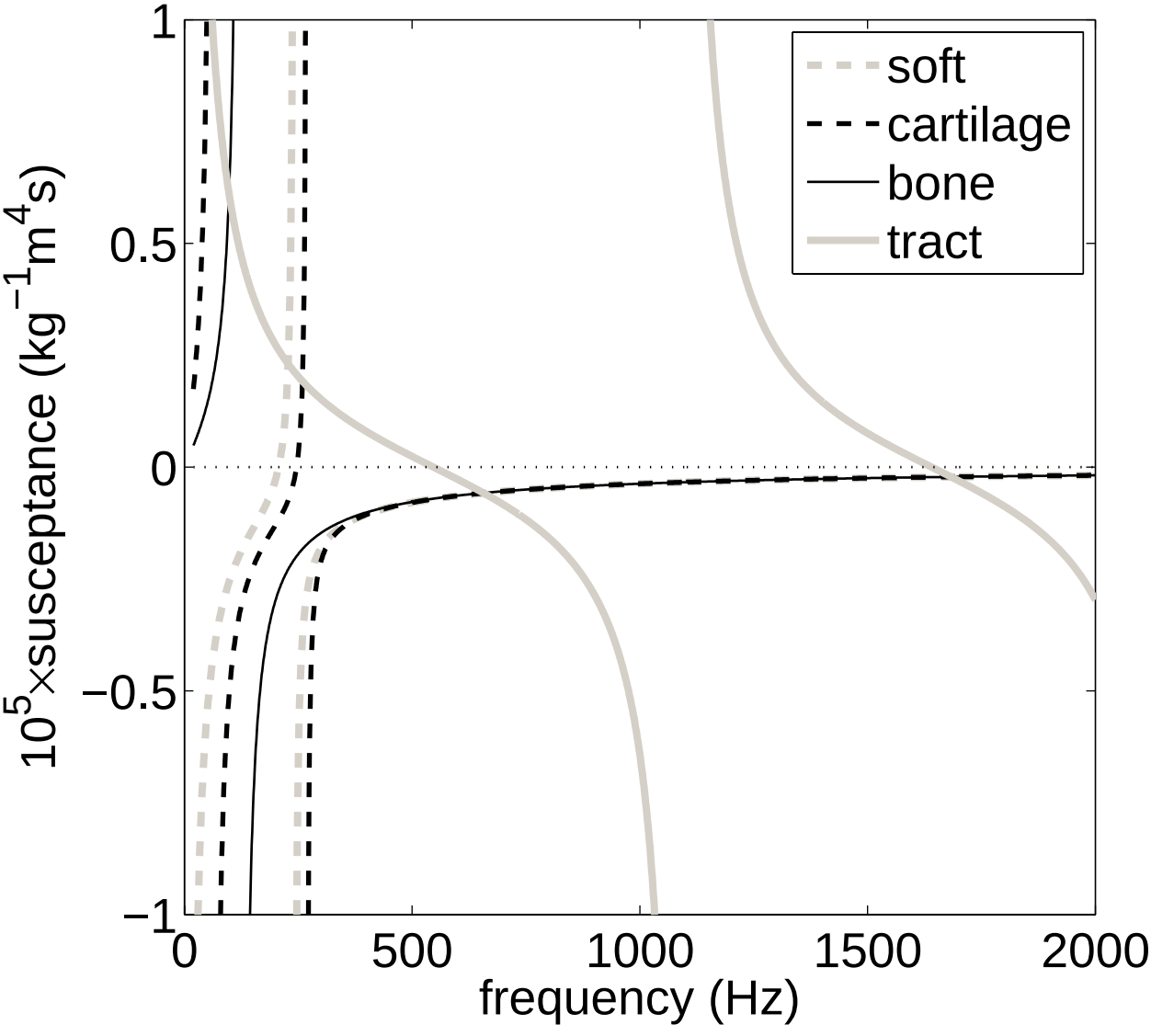

Neck
Cavity
Wall
Radiation
impedance (kg·m⁻⁴s⁻¹)
Frequency (Hz)
ωL_N
R_N
1/ωC_C
R_C
1/ωC_W
ωL_W
R_W
bone
bone
cartilage
soft
cartilage
soft
ωL_r
R_r
total

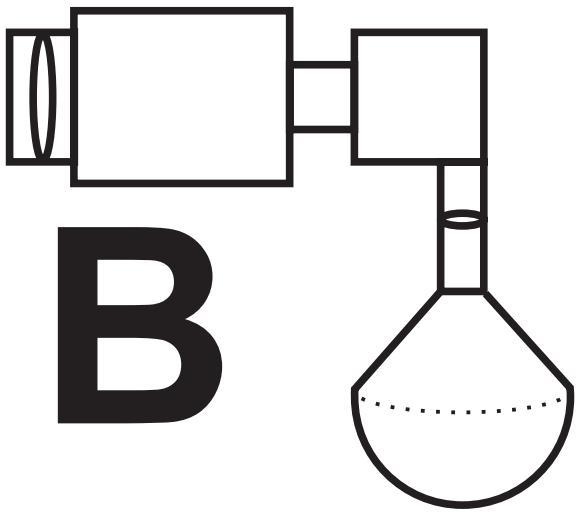

10^{10}
10^{8}
10^{6}
10^{4}
impedance (kg·m^{-4}s^{-1})
20
50
100
200
500
1000
2000
4000
frequency (Hz)
soft
cartilage
bone


soft
cartilage
bone
10^{10}
10^{8}
10^{6}
10^{4}
impedance (kg·m^{-4}s^{-1})
20
50
100
200
500
1000
2000
4000
frequency (Hz)

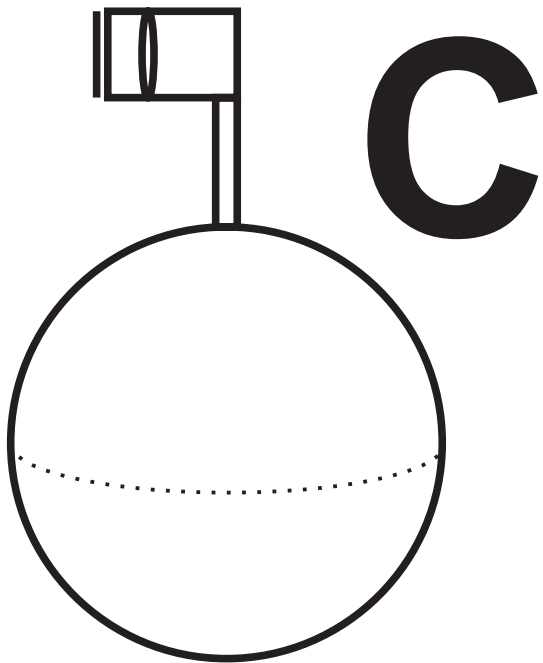

power (W)
10^{3}
10^{0}
10^{-8}
0
500
1000
1500
2000
2500
3000
3500
4000
frequency (Hz)

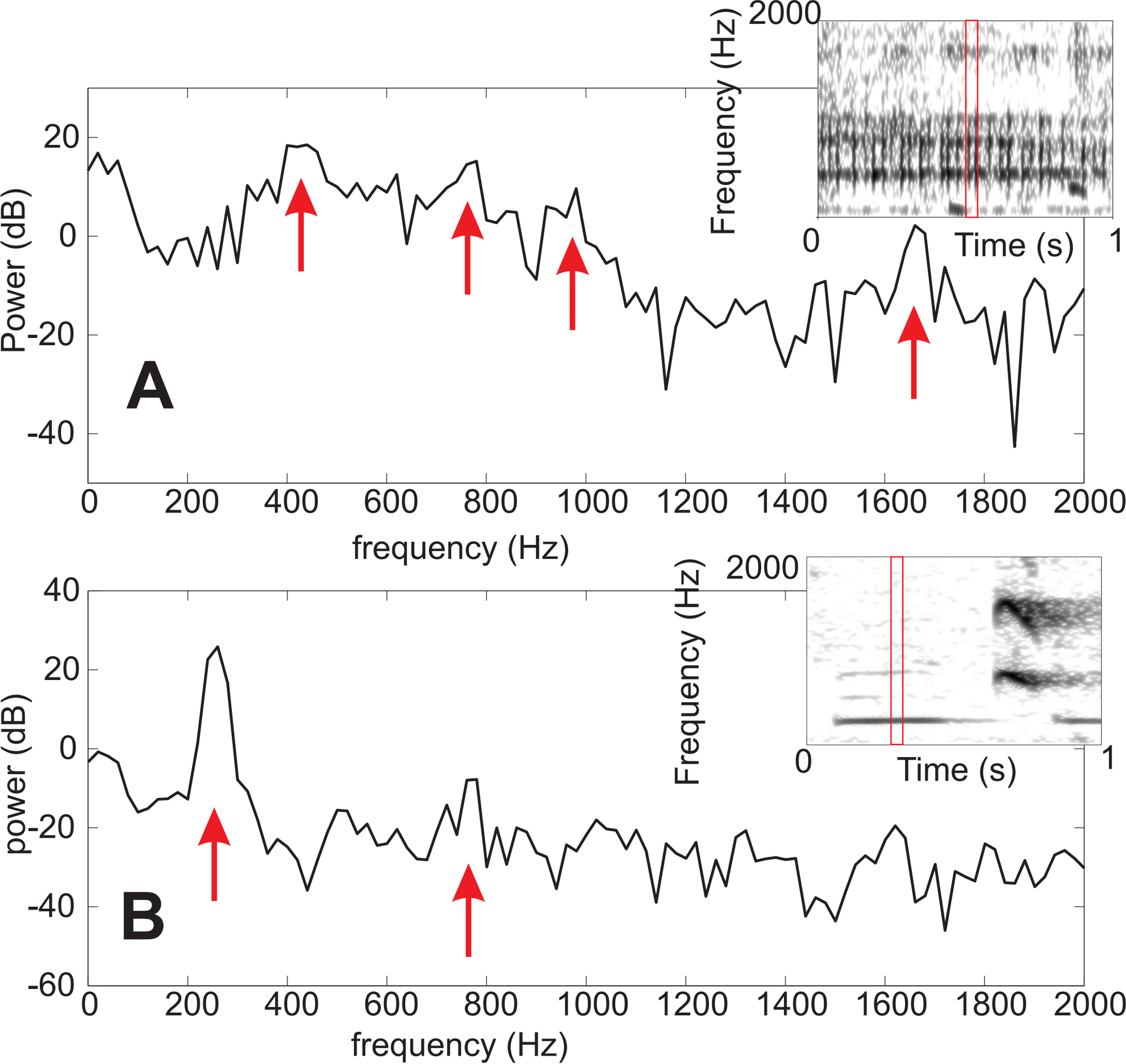

ratio P(sac)/P(mouth)
10^3
10^0
10^−5
10^−10
0
500
1000
1500
2000
2500
3000
3500
4000
soft
cartilage
bone

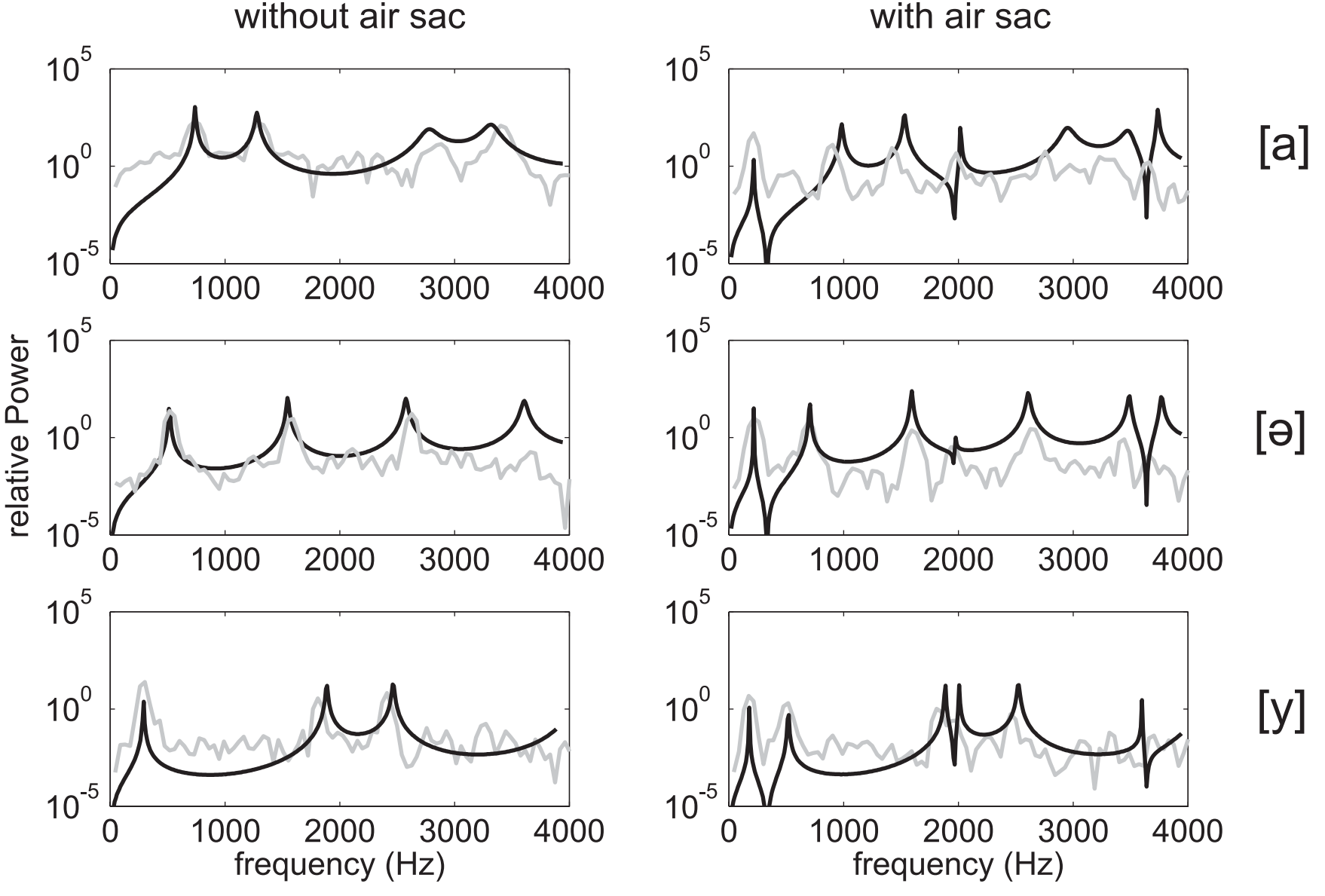

soft
cartilage
bone
tract
1
0.5
0
−0.5
−1
0
500
1000
1500
2000
frequency (Hz)
$10^5 \times$susceptance ($kg^{-1}m^4s$)

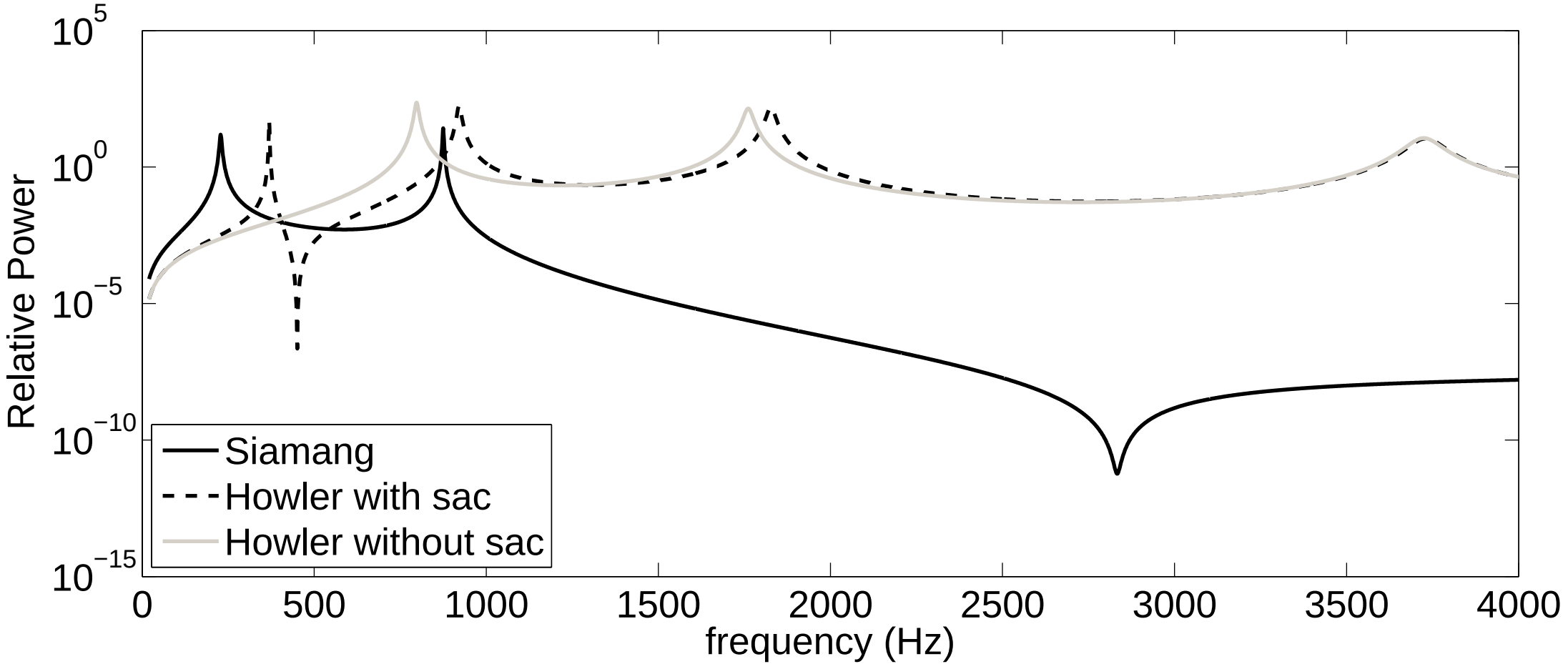


10 cm




A
Power (dB)
20
0
-20
-40
0
200
400
600
800
1000
1200
1400
1600
1800
2000
frequency (Hz)
Frequency (Hz)
2000
0
Time (s)
1
B
power (dB)
40
20
0
-20
-40
-60
0
200
400
600
800
1000
1200
1400
1600
1800
2000
frequency (Hz)
Frequency (Hz)
2000
0
Time (s)
1


without air sac
with air sac
relative Power
frequency (Hz)
frequency (Hz)
[a]
[ə]
[y]
10^5
10^0
10^{-5}
0
1000
2000
3000
4000


Relative Power
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
Siamang
Howler with sac
Howler without sac
0
500
1000
1500
2000
2500
3000
3500
4000
frequency (Hz)