**Supplemental Materials**

**Methods**

*Two mass model of the vocal folds with vocal membrane*

A mathematical model was constructed for simulating the vocal fold oscillation with vocal membranes (Mergell *et al.*, 1999). Each side of the vocal fold tissue was divided into upper and lower portions of masses, *m*1 and *m*2, which are coupled by springs. As the vocal membrane, a massless rigid reed (length: *d*3) is connected to the upper mass with an angle $θ$.

The dimensions of the vocal folds were based on measurements in grasshopper mice and house mice and set as follows. Length: *l* = 500 m, depth of the lower mass: *d*1 = 500 m, depth of the upper mass: *d*2 = 100 m, thickness: ** = 20 m (for grasshopper mouse), ** = 80 m (for house mouse). The size and angle of the vocal membrane were set as *d*3 = 70 m, ** = 0.05**. The sizes of the lower and upper masses were given by $m\_{1}=ρld\_{1}δ,$ $m\_{2}=ρld\_{2}δ$, where  represents the tissue density of lamina propria ($ρ=1.04$ g/cm3). The stiffness values were given by $k\_{1}=ld\_{1}E/δ,$ $k\_{2}=0.1k\_{1}$, $k\_{c}=(5/16)k\_{1}$, where *E* is the Young’s modulus. The collision constants were set as *c*1=3*k*1, *c*2=3*k*2, while the damping constants were set as ri=2miki)1/2 using the damping ratio of .

Symmetric motion between the left and right vocal folds was assumed. Letting *x*1 and *x*2 be the displacements of the lower and upper masses, the equation of motion reads

$m\_{1}\ddot{x}\_{1}+r\_{1}\dot{x}\_{1}+k\_{1}x\_{1}+Θ\left(-a\_{1}\right)\frac{c\_{1}a\_{1}}{2l}+k\_{c}\left(x\_{1}-x\_{2}\right)=ld\_{1}P\_{1}$,

$m\_{2}\ddot{x}\_{2}+r\_{2}\dot{x}\_{2}+k\_{2}x\_{2}+Θ\left(-a\_{2}\right)\frac{c\_{2}a\_{2}}{2l}+k\_{c}\left(x\_{2}-x\_{1}\right)=ld\_{2}P\_{2}$.

Here, *k*c stands for mutual coupling between the two masses. Lower and upper glottal areas are given by *a*i=*a*0i+2*lx*i using the prephonatory areas *a*0i. Glottal area at the upper edges of the vocal membrane is given by $a\_{vl}=a\_{2}-2ld\_{3}\sin(θ)$. *c*i denotes collision force, which is activated during the glottal closure, where the activation function is defined by $Θ\left(x\right)=1 \left(x>1\right), 0 \left(x\leq 0\right)$. Under the assumption that the flow inside the glottis obeys the Bernoulli principle below the narrowest part of the glottis, the pressure that acts on each mass is determined as

$$P\_{1}(x\_{1},x\_{2})\left\{\begin{matrix}P\_{s},&[1,2,6b]\\P\_{s}\left[1-(\frac{a\_{vl}}{a\_{1}})^{2}\right],&[3,4]\\0,&[5,6a,7,8]\end{matrix}\right.$$

$$P\_{2}(x\_{1},x\_{2})\left\{\begin{matrix}P\_{s}+\frac{a\_{2}}{2ld\_{2}\tan(θ)},&[2,6b]\\P\_{s}\left[\left\{1-(\frac{a\_{vl}}{a\_{2}})^{2}\right\}+\frac{d\_{3}}{d\_{2}}\cos(θ-)\frac{d\_{3}a\_{vl}}{d\_{2}a\_{2}}\cos(θ)\right],&[3,4]\\0,&[1,5,6a,7,8]\end{matrix}\right.$$

where $P\_{s}$ represents the subglottal pressure and the nine classifications are made in accordance with the following glottal configurations: (1) $a\_{1}>a\_{2}=a\_{vl}=0$, (2) $a\_{1}>a\_{2}>a\_{vl}=0$, (3) $a\_{1}>a\_{2}>a\_{vl}>0$, (4) $a\_{2}>a\_{1}>a\_{vl}>0$, (5) $a\_{2}>a\_{vl}>a\_{1}>0$, (6a) $a\_{2}>a\_{vl}>a\_{1}=0$, (6b) $a\_{2}>a\_{1}>a\_{vl}=0$, (7) $a\_{2}>a\_{1}=a\_{vl}=0$, (8) $a\_{1}=a\_{2}=a\_{vl}=0$.

The prephonatory glottal areas were set for upper and lower masses as *a*01 = *a*02 = 18000m2. The Young’s modulus was varied from 1.5 kPa to 4 kPa based on measurements in rat vocal folds (Riede et al. 2011).

Our goal was to investigate the effect of different vocal fold morphologies on the minimum lung pressure that would initiate vocal fold oscillation (= phonation threshold pressure, PTP), as well as the relationship between lung pressure and fundamental frequency. To detect the PTP, first, the Young’s modulus was fixed to a constant value and subglottal pressure $P\_{s}$ was increased from 0 to 4 kPa until the vocal folds start to oscillate. At this point, fundamental frequency and PTP were measured to draw the 2-dimensional graph in Fig. 6.

*Vocal tract model*

Acoustical characteristics of the flared vocal tract were estimated using the transmission line model (Flanagan, 1972; Sondhi and Schroeter, 1987). In this model, the vocal tract is represented by a series connection of many short uniform tubes. Assuming a plane-wave propagation, the input-output relationship of the pressure and the volume velocity at each tube is described in terms of a four terminal network. Acoustical characteristics of the vocal tract are obtained by computing the transmission characteristics of the cascaded connection of such lumped elements. As the acoustic impedance parameters, standard values that assume a static wall condition were utilized (Flanagan, 1972). Radiation impedance at the open end of the vocal tract was determined according to Causse et al. (1984).

As show in figure 5, the vocal tract is composed of 7mm-long uniform tube and 13mm-long flared tube. The radius of the uniform tube was set to 1mm. For the flared tube, the input radius was set to 1mm, while the end radius was varied from 1mm to 5mm. In the calculation of the transmission line model, the vocal tract tube was divided into 200 lumped elements.

**References**

Causse R, Kergomard J, Lurton X. 1984 Input impedance of brass musical instruments. *J.* *Acoust. Soc. Am.* **75**, 241-254.

Flanagan JL. 1972 Speech analysis, synthesis and perception. Springer, Berlin.

Mergell P, Fitch WT, Herzel H. 1999 Modeling the role of non-human vocal membranes in phonation. *J. Acoust. Soc. Am.* **105**, 2020–2028.

Riede T, York A, Furst S, Müller R, Seelecke S. 2011 Elasticity and stress relaxation of a very small vocal fold. *J Biomechanics* **44**, 1936-1940.

Sohndi MM, Schroeter J. 1987 A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Trans. Acoustics, speech, and signal processing* **35**, 955-967.