1-Dimensional Model for Vocal Fold Vibration Analysis
BENG 221 Problem Solving Session

10/4/2013
Group 2
Ana L. Bojorquez, Mieko Hirabayashi, John Hermiz, Pengzhi Wang

Table of Contents

Introduction

There are several motives to model different aspects of the voice box and the movement of the vocal chords, among which is the development of better synthetic speech sensors. In this study, we attempt to analytically and numerically approximate the displacement over time of a vocal cord with both a constant force and impulse stimuli. We applied a simplified model of an oscillatory mass and spring system with damping, to account for the properties of the environment of the vocal cords. However, to create a proper model, we must first understand the anatomy of the voice box.

The Vocal Cords: Anatomy

There are two sets of vocal cords, the posterior or ventricular chords, and the inferior set of "true" folds. The primary functions of the ventricular cords are to lubricate the true vocal chords, and to block food and liquid from entering the airway. They are not actively involved in phonation, except for yelling, or grunting [7]. The inferior set of vocal cords are primarily involved in phonation, they are protected by a layer of stratified squamous epithelium, or a mucous membrane, and are attached to the adductor muscles. The adductor muscles close, joining the vocal cords together, and providing resistance to exhaled air from the lungs [3]. The air then rushes through the vocal cords pushing them aside, and the pressure between the cords drops and sucks them together causing a "Bernoulli Effect". This vibration then produces sound which is then shaped by muscular changes in the throat, jaw, tongue, palate, and lips, creating speech.

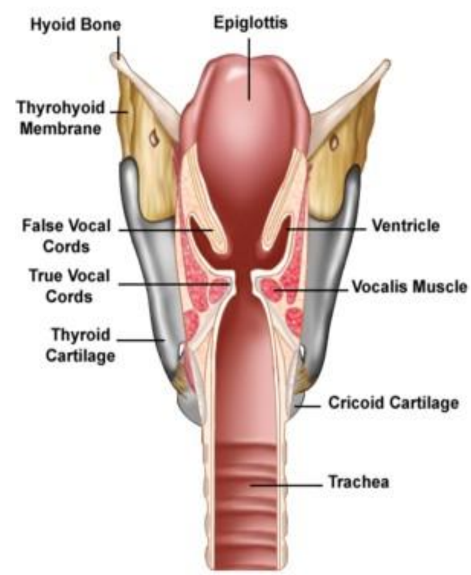


Figure 1: Cross sectional view of Phonetic box. [9]

Injuries to the Vocal Chords

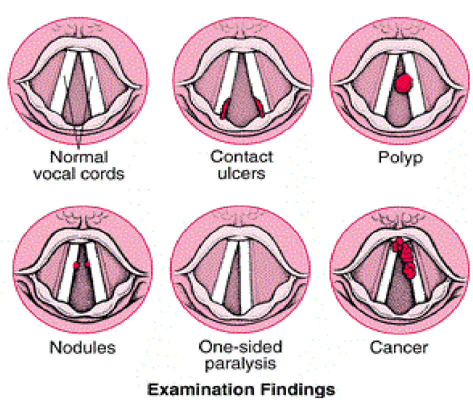


Figure 2 Examples of vocal cord injury. Injuries in this area are difficult to treat due to the overuse and misuse of the vocal chords. [17]

Proper understanding of the physiology of the voice box improves medical practices on the system, and also personal use of the voice. The most common source of injury to the vocal cords is related to improper voice use. Whether it be voice overuse, misuse due to poor singing techniques, coughing, or even acid reflux, any treatment of the vocal cords is difficult due to the fact that they are in constant use.

Some examples of vocal cord injuries, as shown on figure 2, include nodules, granulomas, polyps, contact ulcers, and even cancer tumors. Vocal cord nodules are formed on both sides and are the result of misuse and abuse of the voice like constant yelling, or improper singing [16]. Other injuries like granulomas and polyps can be the result of gastroesophageal reflux, inhalation of irritant chemicals, even endotracheal intubation techniques at a hospital [16]. Thus we can see how any injury in this area would greatly impair the patient's speech abilities and make it very slow for the healing process to occur, so any technological development leading to the expediting of this process is worth investigating.

Current Technology

Most devices used today involve the choosing of words or icons from a computer screen, or tablet, to formulate sentences [12]. This system is also used by the renowned Dr. Stephen Hopkins, whose overall system reads a muscle twitch from his right cheek which Dr. Hopkins uses to choose words, and prepare sentences and speeches in advance. A current study by Gene Ostrovsky, is currently trying to create an artificial larynx which can be programmed to read the interactions of the tongue and the palate during normal speech. This device would then be implanted in the throat of a mute, and produce speech through a speaker box [11].

All of the current technology requires a pre-generated database which is limited by the systems memory capacity, and which requires additional equipment to be carried by the patient. Furthermore, there are limited technologies efficient in treating vocal cord injury alone, not complete impairment. Therefore, with a thorough understanding of the functions and mechanical stresses of the vocal cords, it is possible to develop synthetic materials that can be used to treat vocal cord injury. A current study by a group of students in the Massachusetts Institute of Technology developed a hydrogel composed of hyaluronic acid, and polyethylene glycol-diacrylate which models the human vocal mucosa, and mimics the vibrations of the vocal cords. Gels like these may be further developed to aid in the healing process of injured vocal cords, by removing some of the physical stresses of phonation.

Model

Description

As discussed above, the vocal folds are mucous membranes that vibrate to produce sound [17]. This process is dependent on many variables, including glottal flow, muscle interaction with vocal folds, feedback from the vagus nerve, and material properties of the vocal folds. In order to simplify this simulation, several assumptions and simplifications were made:

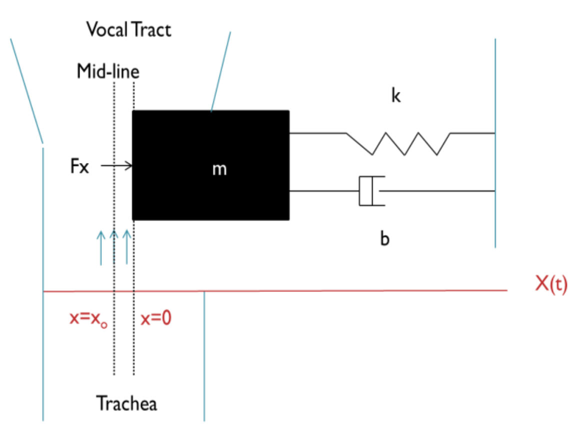


Figure 3: Spring-Mass-Damper Model

(1) Symmetry: since the vocal folds are symmetrical, we assume that we can model one vocal fold independently from the other vocal fold.

(2) One Dimensional: for simplicity, only vibrations in the x direction are being modeled which implies the assumption that vibrations in the y and z direction are independent from vibrations in the x-direction.
(3) Viscoelastic Material: We assume that vocal folds can be modeled as a viscoelastic material with a stiffness *k* and viscosity *b*.
(4) Isotropic Material: we make the assumption that material properties are constant throughout, allowing for material coefficients *k*, *b*, and *m* (mass), to be constant across space (x).
(5) Time Independent Coefficients: due to various types of feedback in the vocal system, material coefficients can be time dependent. For simplicity, we assume that we are modeling for a short enough period of time at coefficients can be approximated as constant over time.
(6) Constant Glottal Flow: for simplicity, we assume glottal flow is constant over a short period of time

With the assumptions made above, the vocal fold model can be simplified as a single mass-damper-spring model with a mass *m*, spring constant *k*, damping coefficient *b*. If the glottal flow is modeled as the forcing function F_x, the equation of motion becomes:

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = F_x$$

Analytical Solution

Because this equation of motion is linear nonhomogeneous differential equation, we decide to solve it analytically by using Laplace transforms. Thus, the equation can be written as shown below with parameters listed in table 1.

$$m \cdot x'' + b \cdot x' + k \cdot x = F_x \quad \text{(i)}$$

Table of Parameters	
Mass (*m*)	0.125-0.4 g [1]
Spring Constant (*k*)	35-80 N/m [2,3]
Damping Coefficient (*b*)	ʒ *2*sqrt(k*m)
Damping Ratio (ʒ)	0-0.2 [2]
Fx	0.01 N

Table 1: The parameters of the single mass-damper-spring model system.

Apply the Laplace operator on both side of the equation (i), which is a linear operator. Then we can get

$$L(F_x)(s) = L(m \cdot x'' + b \cdot x' + k \cdot x)(s) \quad \text{(ii)}$$
$$= mL(x'')(s) + bL(x')(s) + kL(x)(s)$$

According to the table of Laplace transforms, we can further simplify the equation (ii),

$$10^{-4} \cdot [s^2 L_x(s) - s \cdot x(0) - x'(0)] + 0.010955 \cdot [sL_x(s) - x(0)] + 30L_x(s) = 0.1/s$$

with $x(0) = x'(0) = 0$.

By separating the $L_x(s)$ and decomposing the denominator, we get:

$$30L_x(s) + 0.010955 \cdot s \cdot L_x(s) + 10^{-4} \cdot s^2 \cdot L_x(s) = 1/(10s),$$

$$\text{i.e., } L_x(s) = \frac{10^3}{s\cdot(s^2+109.55s+3\times10^5)} = \frac{10^3}{s\cdot(s-s_1)\cdot(s-s_2)} = 10^3 \cdot [\frac{\frac{1}{(s_1-s_2)}}{s(s-s_1)} - \frac{\frac{1}{(s_1-s_2)}}{s(s-s_2)}]$$
$$= \frac{10^3}{(s_1-s_2)} \cdot [\frac{1}{s_1} \cdot (\frac{1}{s-s_1} - \frac{1}{s}) - \frac{1}{s_2} \cdot (\frac{1}{s-s_2} - \frac{1}{s})]$$
$$= \frac{10^3}{s_1(s_1-s_2)} \cdot \frac{1}{s-s_1} - \frac{10^3}{s_2(s_1-s_2)} \cdot \frac{1}{s-s_2} + \frac{10^3}{s_1 \cdot s_2} \cdot \frac{1}{s} \quad \text{(iii)},$$

with $s_1, s_2 = \frac{-109.55 \pm \sqrt{109.55^2 - 4\times3\times10^5}}{2}$.

Then we could apply the inverse Laplace transforms on the both sides of the equation (iii), i.e. $x(t) = {L_s}^{-1}[L_x(s)]$. We could get:

$$x(t) = {L_s}^{-1}[\frac{10^3}{s_1(s_1-s_2)} \cdot \frac{1}{s-s_1} - \frac{10^3}{s_2(s_1-s_2)} \cdot \frac{1}{s-s_2} + \frac{10^3}{s_1 \cdot s_2} \cdot \frac{1}{s}]$$

From the table of Laplace transform, we can get the $x = x(t)$ by plugging in the value of s_1 *and* s_2:

$$x(t) = -6.80449 \times 10^{-5} \times e^{-109.55t} \times sin(536.655t) - 3.333 \times 10^{-4} \times e^{-109.55t} \times cos(536.655t) + 3.333 \times 10^{-4}, t \geq 0.$$

For solving this kind of linear nonhomogeneous differential equation, we can also solve it with undermined coefficients, which method is just like the one in our lecture notes, to find the eigenvalue and generate the general solution, then further determine the particular solution by the method of undetermined coefficients. Here, we just show the analytical solution by using Laplace transform since we have the I.C. After we get the analytical solution, we will directly compare this with the computational solution from ode45 function in Matlab 2012b.

Computational Solution

Analytical vs Computational

The analytical expression is plotted below. The curve increases to its maximum immediately after t=0 seconds. The maximum reached is about 500um, a physiology relevant value [2-5]. After reaching its maximum, the displacement oscillates in an under-damped fashion finally reaching steady state at about 50 ms. Although the model is not completely phenomenology correct, it does exhibit relevant characteristics.

Model Limitations

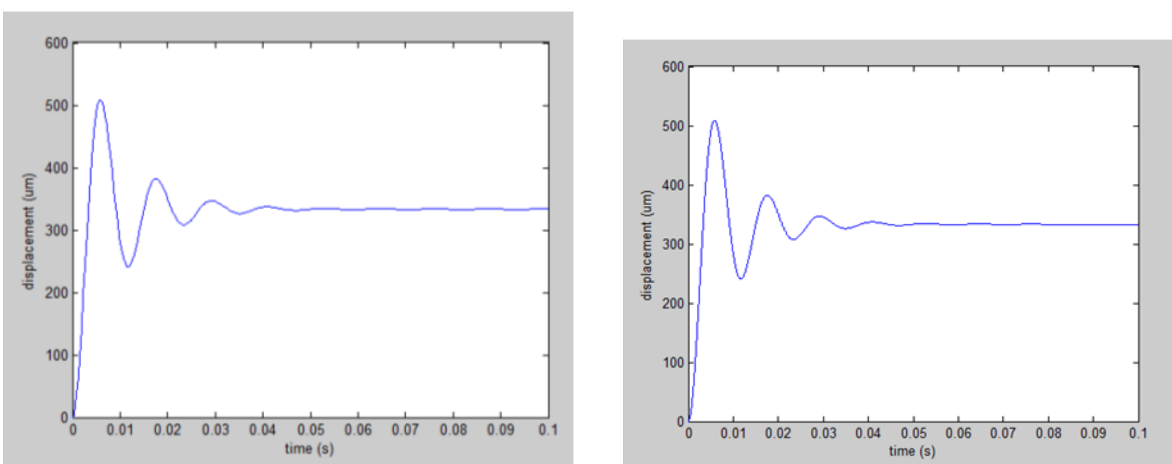
The most significant discrepancy between the analytical model and real vocal folds is that it is an under-damped as opposed to a system that sustains oscillations [2, 22]. However, a system that sustains oscillations, assuming a damper is present in the model, would need an appropriate driving force, F_x. Although setting F_x to a constant was convenient for finding a solution to the model, it does not allow for the interesting dynamics observed in real vocal folds.

Model Relevancies

The model exhibits some characteristics that are applicable to vocal folds. The range of displacement (100's um) and time scale (10 ms) are physiological plausible scales for vocal cords [2-5]. Additionally the model is stable, which is true of vocal folds. The plot of the s-domain below shows that the poles are in the left half plane with the exception of one pole at the origin. In real vocal folds, it is suspected that poles would also be close to the imaginary axis since the vocal folds exhibit self-oscillation.

Computational Plot

Shown alongside the analytical plot, is a numerical solution obtained from the same differential equation. Matlab 2012b software was used to solve the differential equation – specifically, the ode45 function was used to compute the solution. The computational solution appears to be virtually identical with the analytical solution validating the derivation.



Computational

Analytical

Figure 4: Computational and Analytical solution

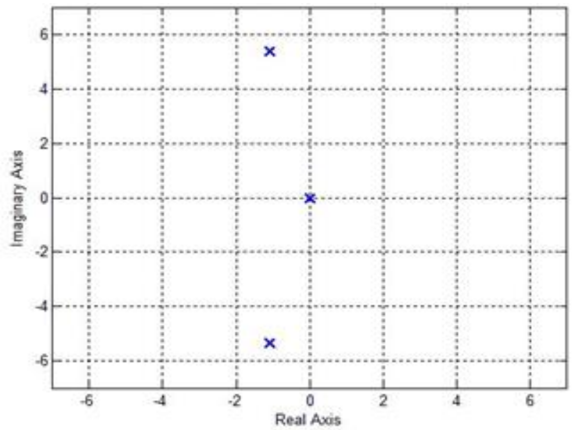


Figure 5: Stability Analysis from S-Domain Plot

Alternative F_x

As mentioned previously, the most significant deficiency in the analytical model was the constant force term, $F_{x.}$ This was chosen mainly for its simplicity and convenience in solving the model. However, with the aid of computers, it is equally convenient to numerically solve for non-linear driving forces.

Therefore, we proposed a new model where the driving force was now proportional to the square of the displacement. This is physically motivated by the idea that as the vocal folds are displaced more air flow is allowed through the glottis, the open region in between vocal folds. This increase is air flow enhances the Bernoulli Effect causing the vocal folds to open more. Eventually this effect is balanced out by the spring (modeling the inherent elastically), which pushes the vocal folds back towards the origin. The damper impedes the Bernoulli force as well.

In addition to the x^2 proportionality, a boundary condition was imposed on the system. Whenever the displacement was such that the vocal fold met with the other vocal fold, there would be an immediate constant force that pulled them a part. This boundary condition is not phenomenology accurate [2]; however, it is a better approximation than was previously made (eg. no condition).

The result of this model is shown below. Once the system is activated (at 5 ms), the displacement is oscillatory, which is comparable to real vocal folds. Although the minimum displacement is greater than the point where it would meet the other vocal fold (-240um), its absolute value is less than the maximum yielding a slightly asymmetric oscillation, which is more characteristic of vocal folds [2].

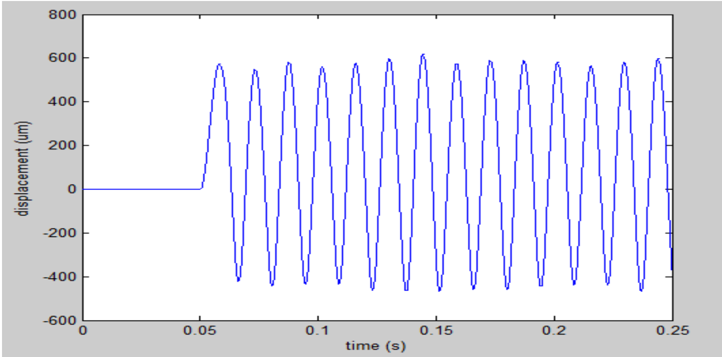


Figure 6 : Computational model from Alternative Fx

Effects of Material Properties on Model Output

Figure 7 in the previous section show one possibility for vocal fold oscillations. However, this behavior changes dramatically when material parameters, k, ʒ, and m, are altered. Figure 8 shows parameters for a theoretical healthy adult. Given various reported values for the mass, spring constant, and damping ratio, average values [1-5], the average values were calculated and found to be m=0.25g, ʒ=0.15, and k=30 N/m. The response of the system shows a maximum displacement of about 500 um after the initial force is applied. This is similar to other models in literature [2]. The response also shows a damped oscillating signal, which is an expected response for any under-damped system (ʒ<1). This is expected because it accounts for the viscoelastic properties of the vocal fold. Since the dominate frequency is not obvious in Figure 8a, a Fourier transform was done on the response to obtain a frequency spectra (shown in Figure 8b). The spectra shows the dominant frequency to be about 335Hz, which is higher than normal speech for an adult male (125 Hz) or female (200 Hz) [24] normal talking voice, but slightly lower than an adult singing voice of A above middle C (440 Hz) [25], indicating that the model gives a reasonable output frequency for a healthy adult.

Children tend to have higher frequency voices than adults. This is partially due to the difference in material properties; children's vocal folds usually have less elastin [24] which leads to the material to have less elasticity and less mass. Therefore, in order to change the model to give a child's response, the values of k and m were both reduced. Figure 6 shows an output that is similar to the healthy adult except the damping of the signal takes longer, which is expected, and the displacement increase. Though the overall shape is the response expected, the increase in displacement does not accurately represent the change from adult to child. This is mostly due to the fact that the mass-damper-spring model does not account for the area of the vocal folds allowing the mass to oscillate past what would be the outer boundary where the vocal folds attach to the throat. While the displacement seems to be modeled inaccurately, the frequency does increase as expected.

Pathologies of the vocal folds also change the material properties of the vocal folds. Most pathologies, such as vocal fold polyps, sore throat, hoarse voice, or larynx reflux, change the sound of an individual's voice due to the inability for the individual to close the vocal folds completely. With a right vocal fold polyp, for example, the mass (polyp) can physically block the closure of the vocal fold, or the increase in mass, stiffness, and damping will require a force much stronger to vibrate the vocal folds [6,7]. In larynx reflux, the acid irritation causes inflammation which also increases the mass, stiffness and damping [7]. The effect of these kinds of pathologies can be seen in Figure 9 and Figure 10, where the displacement is inhibited substantially (about 50%). The Frequency spectra indicate that the input signal simulating the glottal flow is not enough to produce a dominant frequency that would produce sound.

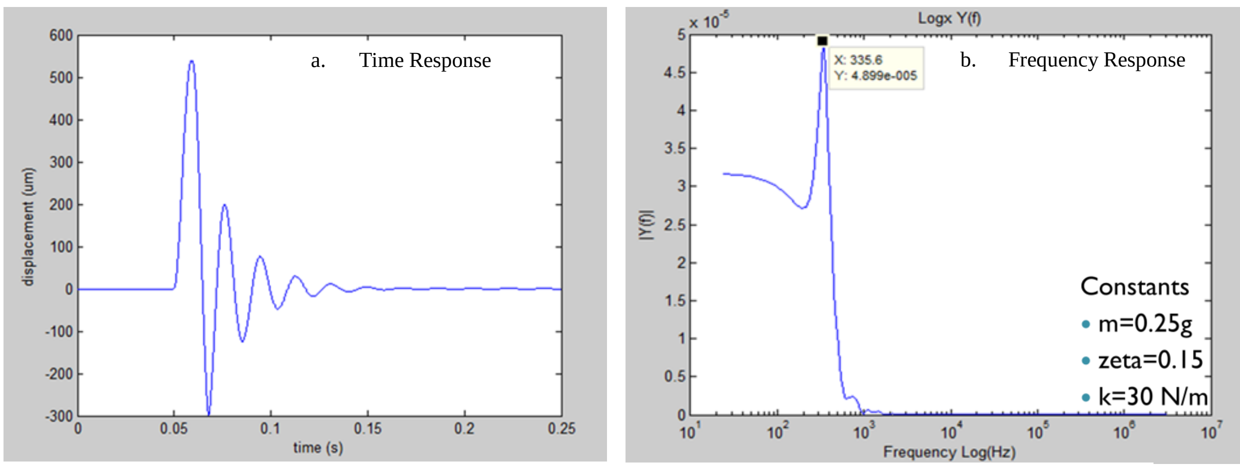


Figure 7: Theoretical Healthy Adult

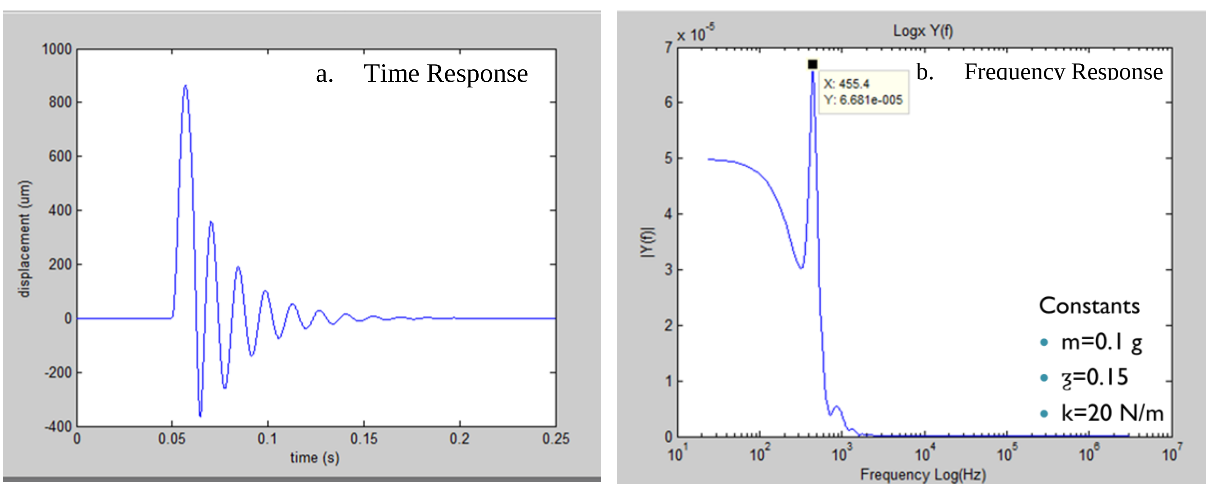


Figure 8: Theoretical Child

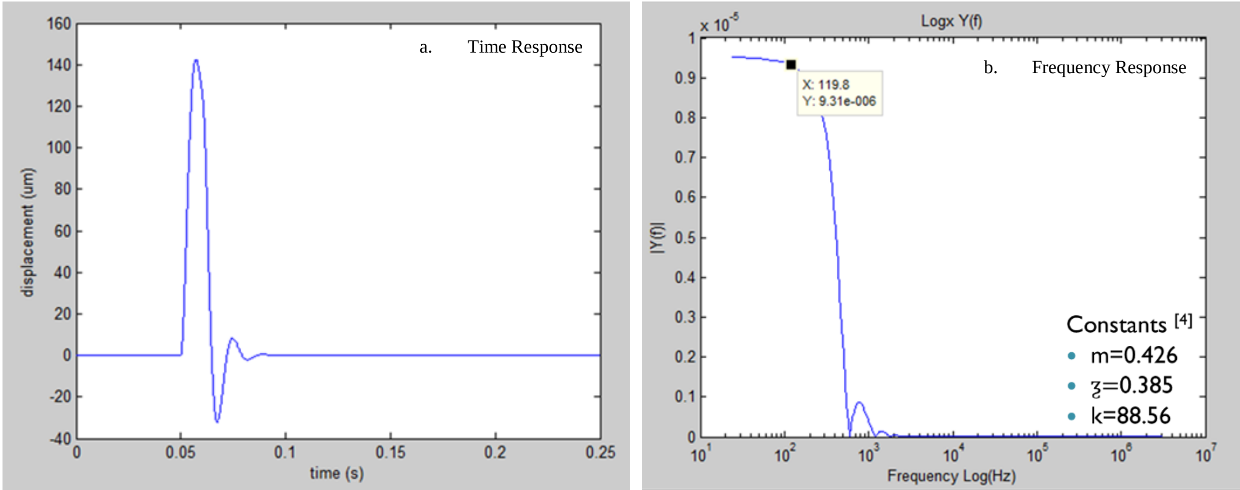


Figure 9: Right Vocal fold Polyp

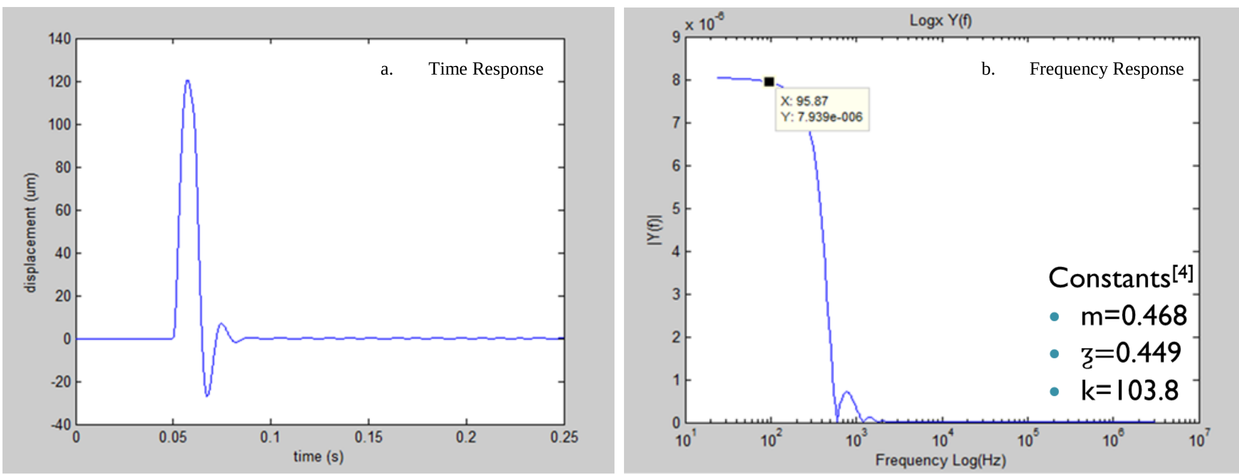


Figure 10: Larnyx Reflux

Conclusions

In order to simply the model of the vocal folds, many assumptions were made. The geometry was lumped which makes the model inaccurate when the geometry in the x-direction changes, such as in the case of the child and adult. Interactions between different tissues and different direction were taken to be independent from the movement of the vocal folds in the x-direction when it is well known that the vocal folds respond to feedback from surrounding tissue and neural input. The glottal flow was also simplified. However, despite all the simplifications, the model still is able to demonstrate how changes in physiology can alter the change in displacement and frequency output.

References

[1] Cataldo, Edson, Christian Soize, Christophe Desceliers, and Rubens Sampaio. "Uncertainties in mechanical models of larynx and vocal tract for voice production." In *Proceedings of the XII International Symposium on Dynamics Problems of Mechanics (DINAME 2007)*. 2007.

[2] Flanagan, J., and Lois Landgraf. "Self-oscillating source for vocal-tract synthesizers." *Audio and Electroacoustics, IEEE Transactions on* 16, no. 1 (1968): 57-64.

[3] Lulich, Steven M. "ESTIMATION OF LUMPED VOCAL FOLD MECHANICAL PROPERTIES FROM NON-INVASIVE MICROPHONE RECORDINGS."

[4] Gómez, P., C. Lázaro, R. Fernández, A. Nieto, J. I. Godino, R. Martínez, F. Díaz et al. "Using biomechanical parameter estimates in voice pathology detection." In *Proc of 4th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications (MABEVA05)*, pp. 29-31. 2005.

[5] Mehta, Daryush D., Matías Zañartu, Thomas F. Quatieri, Dimitar D. Deliyski, and Robert E. Hillman. "Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy." *the Journal of the Acoustical Society of America* 130, no. 6 (2011): 3999.

[6] Thomas, James." Laryngology 101 HD - Hemorrhagic vocal cord polyps, part I, HD" <http://www.youtube.com/watch?v=YFs4etPQd7M>

[7] Thomas, Ja.es "Laryngology 101 - Reflux Laryngitis" <http://www.youtube.com/watch?v=s3NpwHK2REg>

[8] http://www.getbodysmart.com/ap/respiratorysystem/larynx/folds/tutorial.html

[9] http://www.gbmc.org/home_voicecenter.cfm?id=1552

[10] http://en.wikipedia.org/wiki/Speech_synthesis

[11]http://www.medgadget.com/2009/12/artificial_larynx_to_give_mute_a_new_voice.html

[12] http://singularityhub.com/2010/05/03/how-does-stephen-hawking-talk-video/

[13] http://en.wikipedia.org/wiki/Speech_synthesis

[14] http://etd.lib.clemson.edu/documents/1219848273/umi-clemson-1631.pdf

[15] http://www.youtube.com/watch?v=lOJAWOK1RTs

[16]http://www.merckmanuals.com/home/ear_nose_and_throat_disorders/mouth_and_throat_disorders/vocal_cord_polyps_nodules_and_granulomas.html

[17] http://en.wikipedia.org/wiki/Vestibular_fold

[18] http://www.getbodysmart.com/ap/respiratorysystem/larynx/folds/tutorial.html

[19] http://www.gbmc.org/home_voicecenter.cfm?id=1552

[20] http://en.wikipedia.org/wiki/Speech_synthesis

[21]http://www.medgadget.com/2009/12/artificial_larynx_to_give_mute_a_new_voice.html

[22] Fant, Gunnar. *Acoustic theory of speech production*. No. 2. Walter de Gruyter, 1970.

[23] http://en.wikibooks.org/wiki/Speech-Language_Pathology/Stuttering/Core_Behaviors

[24] http://en.wikipedia.org/wiki/Vocal_folds

Appendix A: Matlab Code

Plot Analytical Solution

```
%Plot Analytical Solution
t=0:10e-6:0.1;
x=-6.80449e-5*exp(-109.55*t).*sin(536.655*t)-3.333e-4*exp(-
109.55*t).*cos(536.655*t)+3.333e-4;

plot(t,x*1e6) %convert displacement from m -> um
```

Computational Modeling

```
function main()
clear all
close all

k=30; %spring constant N/m
m=.1e-3; %vocal fold mass kg
zeta=0.1;
b=2*sqrt(k*m)*zeta; %damping factor
l=18e-3; %18mm
d=2e-3; %2mm
Ago=5e-6; %m^2 %glottial area
Ps=8*98; % pascals, pressue due to bernoulli effect when vf's closed
xo=-Ago/l; %threshold where vocal folds meet

df=.01; %amplitude of activation pulse
ic=[0 0]; %initial condition

%differential equations

%constant forcing function
%func= @(t,x) [ x(2) ; (1/m)*(df-b*x(2)-k*x(1))];
%nonlinear forcing function
func= @(t,x) [ x(2) ; (1/m)*(df*pulse(t)+bf(x(1),t)-b*x(2)-k*x(1))];

%solution at forced time points
[t F]= ode45(func,[0:0.000001:0.1] ,ic);

%plot solution
```

```
figure
plot(t,1e6*F(:,1))
xlabel('time (s)')
ylabel('displacement (um)')

%fft of solution
[f,Y]=fft_custom(t,F(:,1));

%activation pulse function
  function y=pulse(n)

    if (n>0.05 && n<.06)
      y=1;
    else
      y=0;
    end

  end

%driving force function
  function y=bf(x,tt)

    if x>xo
      y=k/5*x^2;
    elseif tt>0.051 && x<xo
      y=1/2*Ps*l*d;
    else
      y=k/5*x^2;
    end

  end

end
```

FFT

```
function [f,Y]=fft_custom(t,y)

T=t(2)-t(1);
if ~(t(3)-t(2)==T)
```

```
    disp('error: non-linear time vector, t');
    return;
end
Fs=1/T;
L=length(t);
figure
plot(t,y)
title('y(t)')
xlabel('time (seconds)')
NFFT = 2^nextpow2(L); % Next power of 2 from length of y
Y = fft(y,NFFT)/L;
f = 2*pi*Fs/2*linspace(0,1,NFFT/2+1);
Y= 2*abs(Y(1:NFFT/2+1));
% Plot single-sided amplitude spectrum.
figure
plot(f,Y)
title('Y(f)')
xlabel('Frequency (Hz)')
ylabel('|Y(f)|')

figure
semilogx(f,Y)
title('Logx Y(f)')
xlabel('Frequency Log(Hz)')
ylabel('|Y(f)|')


figure
loglog(f,Y)
title('Logx Logy Y(f)')
xlabel('Frequency Log(Hz)')
ylabel('Log(|Y(f)|')



end
```