Two-Dimensional Physical Modeling of the Human Vocal Tract using Computer-Aided Design

Alexa Julson & Dr. Nathan Miller | Dept. of Physics & Astronomy

Abstract
Precise shaping of the surfaces of the human vocal tract is necessary for the controlled production of sound in speech. Vibration of the vocal folds produces a harmonic series with a fundamental frequency that depends on the tension of the folds. The upper vocal tract creates a natural resonator whose shape determines which frequencies of the waveform are enhanced. Through the construction of 2-D physical models of the upper vocal tract, we are determining how well we can reproduce the changes in the acoustic spectrum corresponding to two different vowels: “i” and “a.”

Introductory Information
Formants
Formants are frequency peaks with high amounts of energy that correspond to vibrations within the resonating tract. The energy differences between formants determine the timbre, or quality of the sound produced. Formants are the foundation of data analysis for this research.

Resonators
Humans have natural resonators that are crucial to speech production. Beginning at the vocal folds, an oscillating frequency is produced from air pressure that is sent from the lungs through the folds. Muscles surrounding the vocal tract control the size of the resonator. Diameter and length are both variable. The oral cavity is also variable of its size, where our soft palates and tongues have movement abilities. The resonator ends at the opening of the mouth.

Objective
Using the above information, the main objective for this research aimed to analyze spectra of vowels and their distinct formants. A comparative representation of each vowel and their respective formants can then be produced. It is important to note that vowels can have multiple pronunciations. Once this data was acquired, our next goal was to design two dimensional models of a human vocal tract using CAD modeling software. With these models, similar data can be acquired and compared to the initial data.

Materials & Method
Initial Data Acquisition
Audacity is an audio recording, editing, and spectral analysis software. We used Audacity to record one of us (Julson) speaking and singing vowels. A spectrum can be analyzed from this recording to pull out the first two formants (see figure 1). Once the first two vowels are determined for each vowel phonetic, a plot of 2nd Formant vs. 1st Formant can be generated (see figure 8 in Results). This process was repeated for 9 phonetics.

CAD Software Design
OnShape, a computer design program, was used for the following process. We were able to design and locally print the models. The models and their corresponding design templates are shown below. Each 3D printed model has the following dimensions: 3 inches x 3 inches x 1 inches D.

Final Data Acquisition
A harmonic sound source was played through a headphone at the opening of the model that corresponded to the vocal cords. We used a microphone to pick up the sound at the opening of the model that corresponded to the mouth opening. We then used Audacity to record the microphone audio and analyze each spectrum.

Data
The following tables hold the collected initial data, and final data. As mentioned in Methods & Materials, the final data was collected using the CAD models.

Initial International Vowel Phonetic Formants (Hz)

<table>
<thead>
<tr>
<th>Formant #</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>150</td>
<td>170</td>
<td>180</td>
<td>190</td>
</tr>
<tr>
<td>Trial 2</td>
<td>160</td>
<td>180</td>
<td>190</td>
<td>200</td>
</tr>
<tr>
<td>Trial 3</td>
<td>170</td>
<td>190</td>
<td>200</td>
<td>210</td>
</tr>
<tr>
<td>Trial 4</td>
<td>180</td>
<td>200</td>
<td>210</td>
<td>220</td>
</tr>
</tbody>
</table>

This data was taken by analyzing the unique spectra for each phonetic listed. Each phonetic is color coded to figure 8 below. The final data sets can be observed in comparison to their corresponding data sets in the initial data.

Note: In the initial data set, trials 1-3 were collected by recording one of our voices (Julson) speaking and singing. Trial 4 was collected using “Pink Trombone”, a vocal sound software.

Results
The following plot is a map of 1st vs. 2nd Formant for each phonetic vowel from the data tables above along with a Stanford University vowel formant diagram that was used as a key for our data plot.

We can determine here that our two final data sets using the 3D printed models agreed with their corresponding initial data. Observe the brown data set in perfect range with the initial “i” (bean) data. Observe the pink data set in perfect range with the initial “a” (barn) data.

Conclusion
Our main goal for this research was to explore the formants in the unique spectra that are produced from vocal vowel production. Our initial data shows a strong relationship between a typical formant frequency range plot (see figure 7) and our acquired data (see figure 8). Once this data was concluded to have an accurate and strong relationship, we were able to obtain CAD models and test their design using the same method of data acquisition. Finally, we can see that our designed models ultimately produced the results that we were expecting.

Acknowledgements & References

Mr. Turner Howard
Senior Instrumentation Specialist – Dept. of Physics & Astronomy
We owe a special thank you to Mr. Howard for guiding the CAD modelling and printing process and for being a tremendous resource and source of support.