



Welcome to Vibrationdata

Acoustics • Shock • Vibration • Signal Processing

February 2004 Newsletter

Greetings

Speech is perhaps the most important characteristic that distinguishes humans from animals.

Speech can take many forms and can serve myriad purposes. Martin Luther King's "I Have a Dream" speech is among the most ennobling in American history.

King's declaration resonated with moral clarity, enhanced by the resonance of his voice.

The purpose of this month's newsletter is to present an acoustical analysis of King's speech, which is given in the second article.

A proper evaluation requires a framework of acoustical principles, which are given in the first article.

Sincerely,



Tom Irvine
Email: tomirvine@aol.com

Feature Articles



An Introduction to Human Speech page 2

**Acoustic Analysis of Martin Luther King's
"I Have a Dream" Speech** page 5

An Introduction to Human Speech

by Tom Irvine

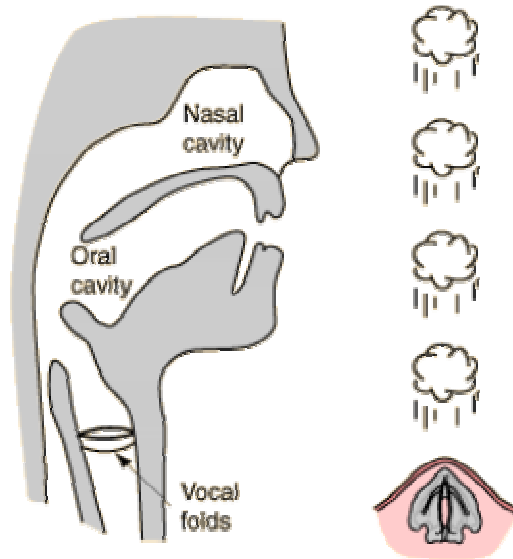


Figure 1. Vocal Folds and Oral Cavity

Image courtesy of HyperPhysics, Georgia State University

Introduction

Speech generation is a rather complex process. This article considers four phases in speech production: respiration, phonation, resonance, and articulation.

Respiration

The lungs provide the airflow through the glottis, which is the opening between the vocal folds. The vocal folds are inside the larynx. The glottis is open during normal breathing. The vocal folds are spread far apart during this phase.

Phonation

Phonation is the process whereby the vocal folds convert the airflow energy into audible sound. Vocal folds are also

referred to as either cords or chords. The folds are muscles.

Additional muscles and cartilages inside the larynx cause the vocal folds to move inward during the onset of speech, closing the glottis. This closing is aided by an aerodynamic effect called the Bernoulli principle. That is, as the speed of a moving fluid or gas increases, the pressure within the fluid decreases. A suction force thus brings the vocal folds together as the airflow moves upward from the lungs to the mouth.

The air particles that have passed through the now closed glottis continue traveling toward the mouth. The remaining pressure in the lungs is thus greater than the pressure on the closed side of the glottis. At a certain pressure

differential, the vocal folds are blown outward, thus opening the glottis and releasing a single 'puff' of air. The elastic restoring force in the vocal folds contributes to this opening process.

The cycle is repeated, producing a periodic train of air pulses, illustrated in Figure 1. The vocal folds control the rate at which this oscillation occurs. Specifically, the vocal folds have a fundamental frequency that is a function of the folds' mass and tension.

The resulting pressure time history would have the shape of a sawtooth wave. This wave is composed of the fundamental frequency and its integer harmonics. It represents the hundreds of air puffs per second that make up speech.

The phonation process thus described is a simplification. Researchers have determined that the folds alternately take on a convergent and divergent shape during the cycle, as shown in Figure 2.

The average air pressure within the glottis tends to be larger in the convergent configuration than in the divergent shape, resulting in the asymmetry of air pressures that helps sustain the oscillation.

Vocal Fold Fundamental Frequency

The most obvious difference between the male and female voice is fundamental frequency, or pitch. Due to the increase in mass of a male's vocal folds, which occurs during puberty, the average speaking fundamental frequency for males varies between 100-132 Hz while the average for females varies between 142-256 Hz, per Reference 1.

Nearly all information in speech is in the range 200 Hz to 8 kHz. Some telephone systems carry sound from only 300 Hz to 3 kHz, but the speech is still reasonably intelligible.

The pitch is determined by the spacing of harmonics perhaps more than by the fundamental frequency, per Reference 2. Thus a man's voice on the phone is readily identifiable even though the fundamental of that signal is not present.

Resonation

Resonation refers to the quality of the voice as regulated by the vocal tract including the soft palate.

The vocal tract is like a closed-open pipe. The natural frequencies of a closed-open pipe of 17 centimeters occur around 500, 1500, and 2500 Hz. The fundamental frequency increases to 600 Hz for if the length decreases to 14 centimeters.

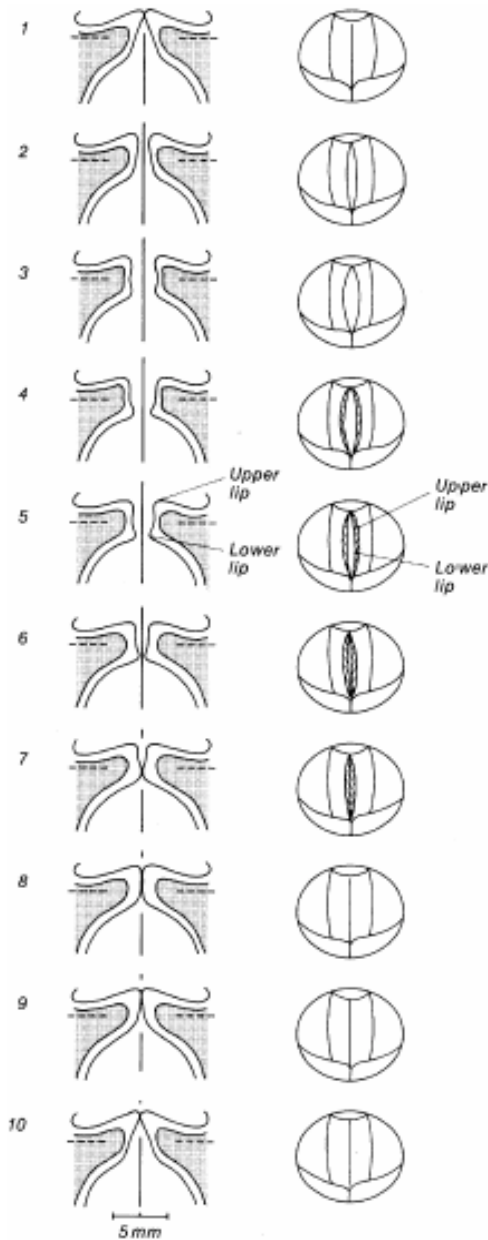
The frequency of speed is determined primarily by the vocal cords. The vocal tract frequency response further shapes the speech production. It acts as a filter that amplifies certain frequencies while attenuating others.

Amplification occurs when the natural frequency of the vocal folds is at or near the natural frequency of the vocal tract. This condition is resonance.

Articulation

Articulators transform the sound into intelligible speech. Articulation is controlled by the positions of the tongue, lips, and jaw. The teeth also play a role.

These articulators retune the natural frequency of the vocal tract system, which is important for producing vowels.



The position of the lips and tongue determine the geometry of the opening, thus controlling the natural frequency of the system.

In this sense, the vocal tract system behaves as Helmholtz or cavity resonator, in addition to behaving as a closed-open pipe.

References

1. Mikos and Pausewang, The Relative Contribution of Speaking Fundamental Frequency and Formant Frequencies to Gender Identification, Presented at the 2001 Convention of the American Speech-Language-Hearing Association November 15-18, 2001, New Orleans, LA.
2. Joe Wolfe, University of New South Wales, 1999.

Figure 2.

Image Courtesy of Phil Hoole



Acoustic Analysis of Martin Luther King's "I Have a Dream" Speech

by Tom Irvine

Introduction

Martin Luther King, Jr. delivered his "I have a Dream" speech on the steps of the Lincoln Memorial in Washington D.C. on August 28, 1963.

King began the speech:

I am happy to join with you today in what will go down in history as the greatest demonstration for freedom in the history of our nation.

Five score years ago, a great American, in whose symbolic shadow we stand signed the Emancipation Proclamation.

King delivered the following memorable lines in the middle of the speech:

I have a dream that my four little children will one day live in a nation

where they will not be judged by the color of their skin but by the content of their character. I have a *dream* today!

King concluded:

When we let freedom ring, when we let it ring from every village and every hamlet, from every state and every city, we will be able to speed up that day when all of God's children, black men and white men, Jews and Gentiles, Protestants and Catholics, will be able to join hands and sing in the words of the old Negro spiritual, "Free at last! Free at last! Thank God Almighty, we are free at last!"

These words deliver a powerful message even in written form. King's pitch modulation and vocal tract resonance transformed his dream into an electrifying elocution which awakened the conscience of a nation.

Vocal Fold Fundamental Frequency

The pressure time history of the “I have a dream” quote is given in Figure 1.

There is a 2 second pause near the middle of the sample. This marks the gap between “children” and “will.”

King exercised tremendous pitch modulation during his speech. This is one of several vocal characteristics that enhanced his message.

Identifying a precise fundamental frequency, however, is challenging as a result of the modulation.

A spectral magnitude function of the time history is given in Figure 2. The magnitude has a linear scale, although the pressure unit is not specified.

The spectral function has sharp peaks at approximately 240 Hz and 360 Hz. The difference between these peaks is 120 Hz. King’s vocal fold fundamental frequency thus appears to be 120 Hz.

Recall from the previous article that the average speaking fundamental frequency for males varies between 100-132 Hz.

The spectral function has a small peak at 120 Hz, which may have been attenuated by the highpass filtering characteristics of the recording equipment. The frequency response characteristics of the recording equipment are unknown, however.

Furthermore, there are numerous spectral peaks across the entire domain in Figure 2. Some of the peaks are harmonics of the vocal fold fundamental frequency. The array of pitches gives a very rich, melodious sound.

Vocal Tract Fundamental Frequency

The highest levels in the spectral function occur among a cluster of peaks from 540 Hz to 650 Hz.

The fifth natural frequency of King’s voice was approximately 600 Hz, which is in the midst of this cluster.

That the fifth natural frequency would project more energy than any of the preceding four would be highly unusual for any system, however.

Vocal tract resonance is the explanation for the cluster response. Recall that the vocal tract behaves as a closed-open pipe. The fundamental frequency of a 14.2 cm long closed-open pipe is 600 Hz. This appears to have been King’s vocal tract fundamental frequency, approximately.

Thus King’s fifth natural frequency excited his vocal tract mode, resulting in significant amplification of his voice.

Pitch Modulation

A spectrogram waterfall plot of the “I have a dream” quote is given in Figure 3. This format reveals the pitch modulation as the spectral peaks shift higher or lower in frequency.

Note the rapid pitch increase that occurs from 540 Hz to 600 Hz from 6 to 8 seconds.

Again, there is a 2 second pause near the middle of the sample, which marks the gap between “children” and “will.”

The frequency increases to nearly 650 Hz as King resumes, pronouncing “will”. Thereafter, the pitch experiences a very gradual decrease, returning to 540 Hz.

The difference between 540 Hz and 650 Hz is approximately one-quarter of an octave, which is a wide spectrum. This domain covers the musical notes C#, D, D#, and E.

Frequency Analysis of a Brief Segment

A time history with a 50 millisecond segment is given in Figure 4. This occurs as the pitch is rising before the gap.

The top signal is the measured data. The bottom signal was synthesized from discrete sinusoids using the method in Reference 1. The goal was to match the characteristics of the measured data. This method indirectly yields the frequencies of the measured data.

The dominant frequency is 560 Hz, which is the modulated vocal tract fundamental frequency.

The signal also contains integer harmonics of this frequency at 1120, 2240, and 3360 Hz.

In addition, there is an 840 Hz component. The mechanism of this sinusoid is not immediately clear, however.

Conclusion

Martin Luther King, Jr. illuminated his call for freedom and justice with his lyrical voice, accentuating his words with pitch modulation and magnifying his message through vocal tract resonance.

Reference

1. Irvine, A Time Domain, Curve-Fitting Method for Accelerometer Data Analysis, AIAA Paper 7667, 2003.

TIME HISTORY - MLK SPEECH EXCERPT

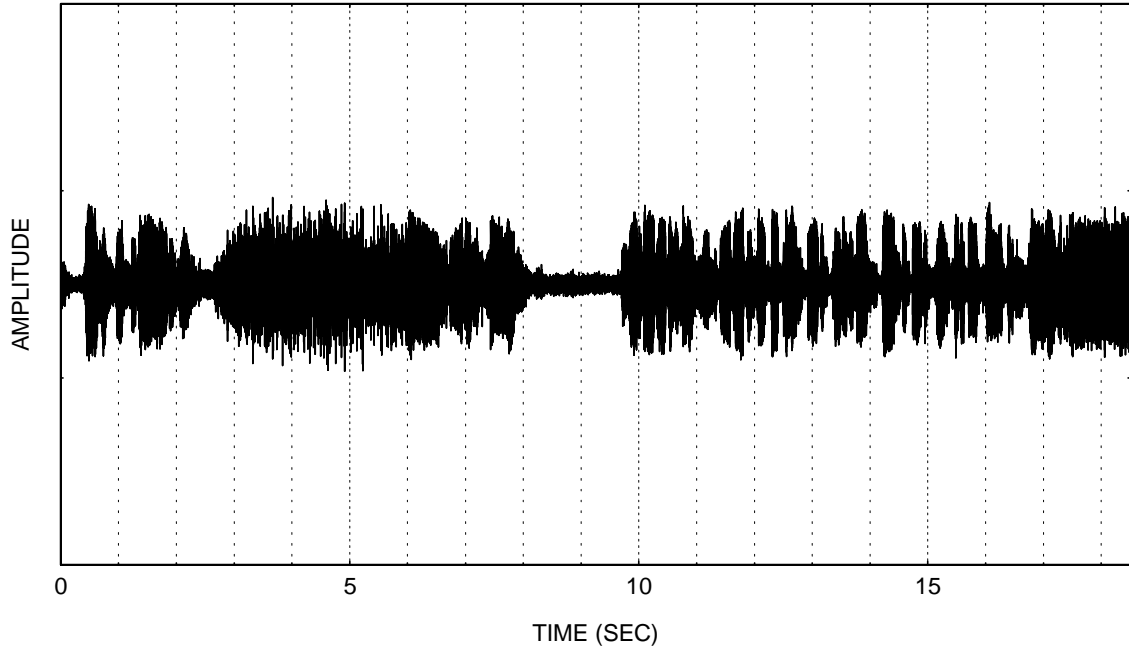


Figure 1.

SPECTRAL MAGNITUDE - MLK SPEECH EXCERPT

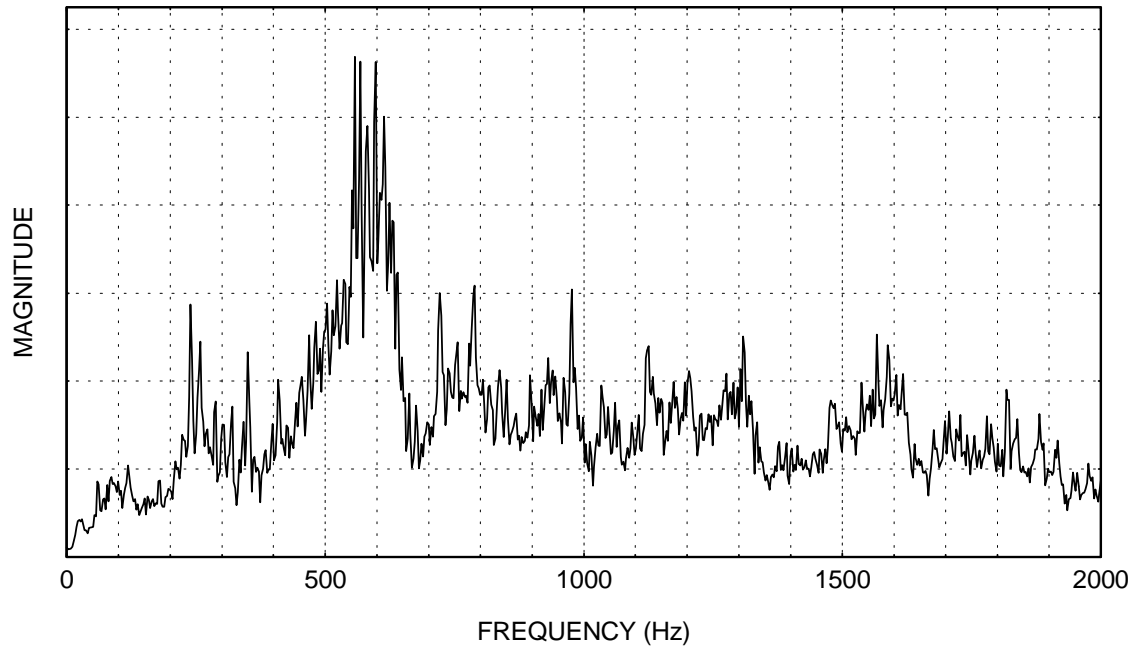


Figure 2.

Spectrogram Waterfall

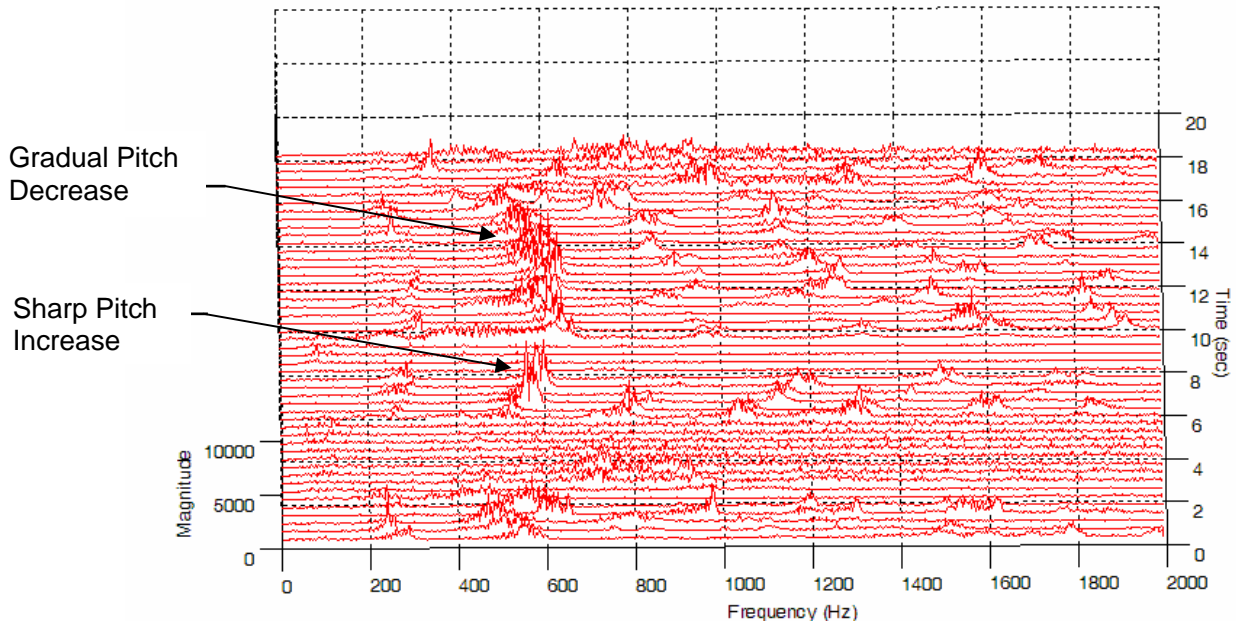


Figure 3.

EXCERPT FROM SPEECH
TOP - MEASURED DATA BOTTOM - SYNTHESIZED SIGNAL

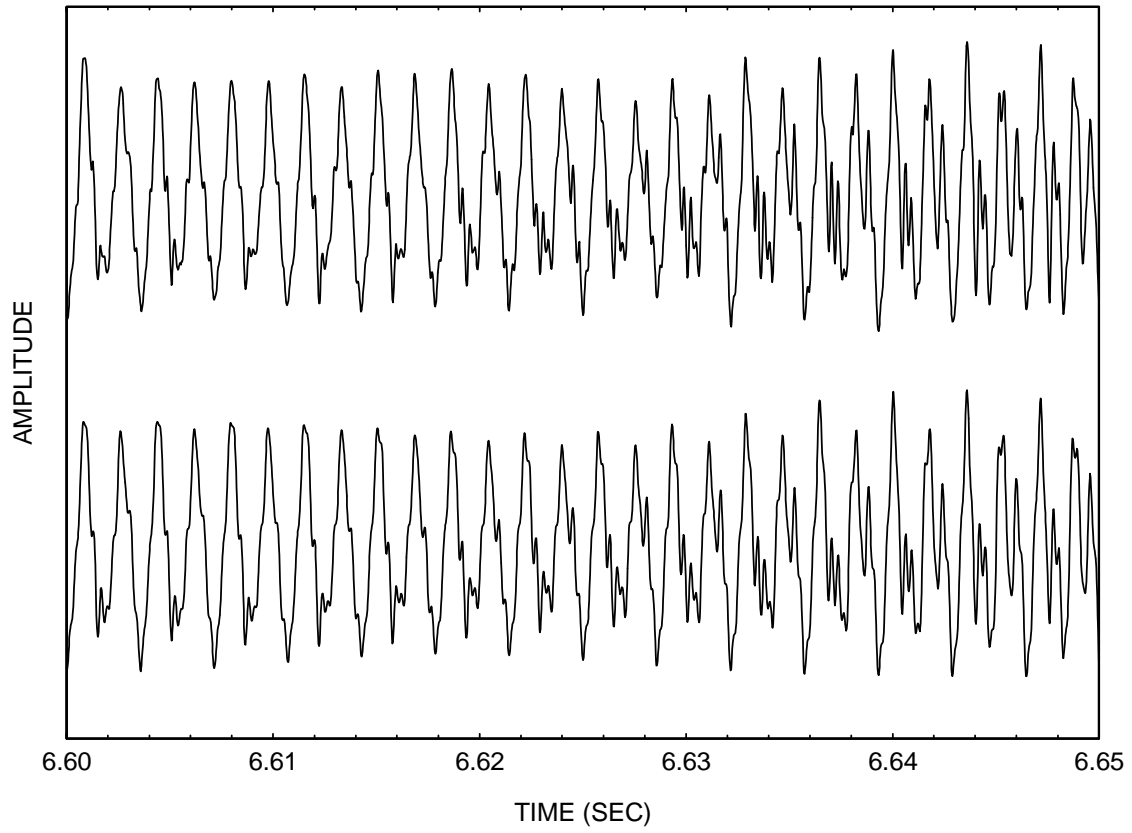


Figure 4.