Wave, resonance, and vowel formants: Application of physics software to education of language and speech science

Yuki Kakita
Kanazawa Institute of Technology, Kanazawa, Japan
E-mail: kakita@his.kanazawa-it.ac.jp

ABSTRACT

Physical basics of speech production, such as sound generation, wave propagation, resonance, standing wave, and vowel formants, are often difficult to understand, particularly for those who are not familiar with physics. This paper proposes that simple physics software can be used effectively to help foster understanding of basic physical and acoustic concepts in speech production. Specifically, basic physical concepts of vowel production are explained by using Interactive Physics, software for modeling, calculating and demonstrating motions based on simple drawings of masses, springs, etc.

1. INTRODUCTION

When we talk about vowel formants in relation to sound spectrograms, for example, we cannot avoid mentioning the resonance of sound in the vocal tract. Of course, we do not need to know about standing waves or resonance if we simply want to relate a vowel directly to a set of formant frequency values. However, if we want to explain that a high first formant in children results from the short length of the vocal tract, we need to know the physical relations between the tract size, the wavelength, and the frequency. This paper proposes that the use of easy-to-use physics education software helps beginners of speech science understand the basic notions of acoustical physics with less pain and more fun. The demonstration of physical phenomena and the illustration of functions should be as realistic as possible even for those who are not familiar with physics and mathematics. For example, the shape of a sinusoidal wave function found in textbooks often has a more rounded shape than the actual function, particularly if the function is hand-drawn based on author's impression. The shapes of functions derived from a precise calculation, instead, is more accurate and more helpful.

2. METHOD

2.1 INTERACTIVE PHYSICS

Interactive Physics [1] is software for modeling, calculating and demonstrating motions based on simple drawings of masses, springs, etc. This paper attempts to show that Interactive Physics can be used to simulate various physical systems with ease and simplicity. The software enables us to demonstrate the behavior of a mechanical system by displaying the graphical output of time varying motions (positions, velocities, accelerations) and forces, etc., based on the graphical input illustrating the structure of the system in the form of, e.g., building blocks.

Interactive Physics can solve simultaneous differential equations accurately for various mechanical systems by employing a number of authorized numerical methods. We can select the level of computational accuracy and speed depending on the problem to be solved. Also, we can view the graphical output of the behavior of the system as calculation progresses. There is no need to wait until the computation is complete. Users basically do not have to worry about numerical methods. If there is any significant problem, e.g., a solution does not converge, Interactive Physics will warn the user with a message, and the user can then set an appropriate value from the menu and restart the computation. There is no actual damage to the simulated system since it is the result of software computation. So the user can try what she/he has in mind without worrying about making mistakes.

After completing a calculation, or stopping a calculation at a certain point, the entire result can be played as an animation. The result, including all the values and animations, can be stored in a computer harddisk for later demonstration and examination. The animation can also be stored as a QuickTime [*] movie, so the result can be played without the Interactive Physics application.

2.2 RELATIVE RELATIONS

In this paper, all demonstrations are performed without specifying the actual values of masses, springs, the speed of waves, or the length of a tract, etc.

When we want to simulate the resonance "modes" of sound as a result of wave propagation in an acoustic tube that approximates a vocal tract, we do not have to specify the actual values for masses and spring constants. The physical reality of the "modes" in vowel production is the "formants". Therefore, what we do with Interactive Physics is simply connect the masses and springs in an alternating series, regardless of the actual values of masses and springs.

 ISBN 1-876346-48-5 © 2003 UAB

Of course we need the actual values when we want to know the vibrating frequency of coupled mass and spring. However, when we refer to standing waves, resonance, and vibrating modes (formants), relations among them are relative and actual values do not matter.

3. RESULTS

3.1 SOUND GENERATION

Sound is defined as the propagation of a minute local displacement of a matter, a group of air particles if the sound is traveling through the air, driven by a mechanical force. In the model used in this paper, air particles — the sound propagation media — are simulated as a series of mass-spring units.

Why is a group of air particles approximated by masses and springs? First, it is quite natural to employ a mass since a group of air particles has its own mass. As for the spring, a group of air molecules in a container is compressible and extensible, exhibiting the spring-like behavior. Consequently, air as the sound propagating medium can be modeled as spring-mass combinations (Figure 1 a, b).

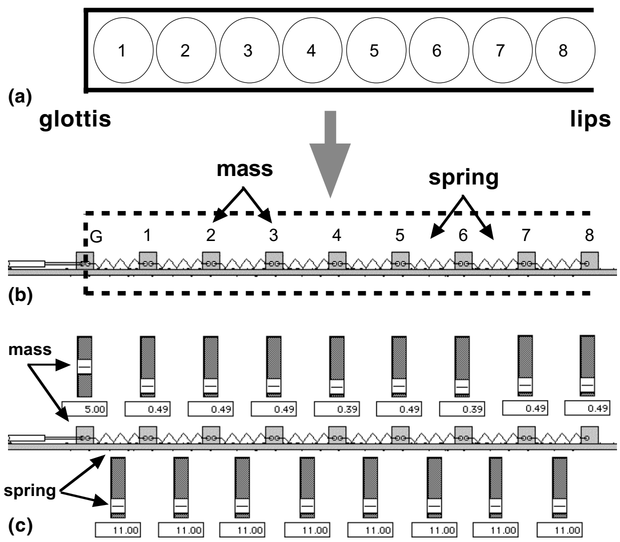

Figure 1 Modeling the vocal tract with Interactive Physics
(a) Groups of air particles as wave propagation medium
(b) Each group in (a) corresponds to a mass-spring unit
(c) Sliders for changing the values of mass and stiffness

In this figure, mass-spring pairs are connected serially, because we generally assume a one-dimensional wave when we discuss resonance in the vocal tract. In other words, in the approximation of a one-dimensional wave, the glottal wave propagates only along the vocal tract, and there is no propagation, for example, perpendicular to the tract wall.

Figure 1 (a, b) illustrates a uniform tube that has one end closed and the other end open, with a uniform cross-sectional area in between. Figure 1 (c), shows a set of sliders each of which represents a mass (upper row) or a spring (lower row). As will be mentioned in later sections, we can create a desired shape of the vocal tract by adjusting these sliders. The number of sections of the tract is limited to 8 in the present model, but the number can be increased depending on computational power.

In the simulations in the present study, driving force is simulated by a velocity source, located at the glottis. (Figure 1 b, c) The velocity source analogy is based on the volume velocity source for the glottis. [2] Strictly speaking, the velocity of a mass is not equal to the volume velocity of a sound. However, in the model used in this paper, a mass does not represent a single air particle but a group (volume) of air particles.

3.2 WAVE PROPAGATION

Sound waves are longitudinal waves and, unlike transverse waves, they are very difficult to visualize or understand. In longitudinal waves, air particles vibrate in the same direction as wave propagation. In the present model, the motion of the mass at one end is shown to travel to the other end, demonstrating, clearly and precisely, the “longitudinal” wave propagation.

Figure 2 shows the wave propagation in a one-dimensional medium. Excitation, a sinusoidal change in one cycle, is applied at the glottis (“0” in the figure). The motion of masses, or the change in velocity representing the “volume velocity of sound”, are shown in rows, each identified by the number. That is, each mass numbered from 1 to 8 (left to right in Figure 2(a)) in Figure 2 corresponds to each row numbered from 1 to 8 in Figure 2 (b). The wave proceeds rightward with time. The downward arrow indicates the transfer of the shape of the source wave. The speed of transfer of the waveform is defined as the “speed of sound”. The “speed of sound”, means the speed of the “shape” of the wave but not the actual movement of a particular particle. We call this kind of speed “phase velocity”.

Figure 2 also demonstrates the reflection of waves, as indicated by upward arrows. The wave is reflected at the open end (mouth).

The wave “reflects” where there is a “discontinuity”, i.e., the difference in the area. The model shown in Figure 2 is that of a uniform tube. Since, in Figure 2, the wave travels through a vocal tract with a uniform cross-sectional area, there is a “discontinuity” at each end.

ISBN 1-876346-48-5

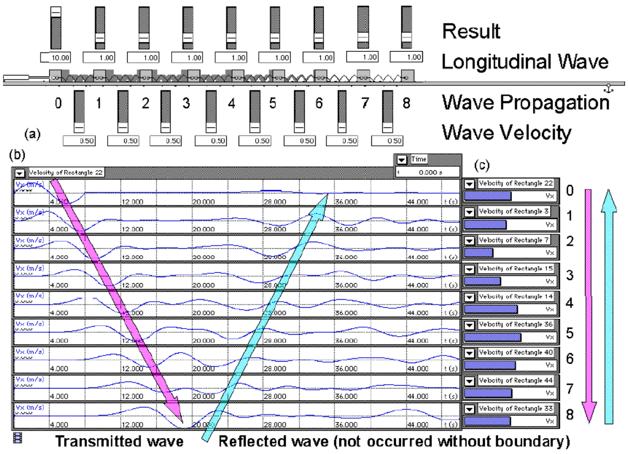

Figure 2 Simulated result of wave generation, transmission, and reflection. The curve in each row (b) indicates the velocity of each mass. Horizontal bar graphs (c) show the values of velocity of the corresponding masses.

3.3 VOCAL TRACT

A "vocal tract" is modeled as a series of mass-spring units, each of which corresponds to a portion of the tract. One end that corresponds to the glottis is terminated by glottal excitation source, or the vocal folds, while the other end corresponding to the mouth is left free, simulating the open air.

Strictly speaking, terminal compensation is necessary in simulating the effect of orifice to open air. For simplicity, however, compensation is not explicitly applied to the present model, assuming that the basic behavior is not affected greatly.

A constriction in the vocal tract is simulated by two heavy masses joined by a rigid spring. Acousto-mechanical relations indicates that the narrower the constriction the heavier the mass and the stiffer the spring constant. [3]

3.4 RESONANCE

The "resonance" is defined by a vibration of a system to be induced by other sources of vibration. The "standing wave" in the vocal tract occurs when the resonance occurs at each mass-spring unit and, at the same time, when these units are connected each other to transfer "motional information" as a shape of "wave".

Generating process of the "standing wave" can be clearly visualized as vibrations of relevant masses at specific driving frequencies. (Figure 3) The transmission and reflection of waves are added together to create standing waves.

Figure 3 shows the resonance of a "uniform acoustic tube" as a result of employing the "sweep tone" measurement. A sweep-tone is a sinusoidal wave whose frequency increases of decreases gradually.

A "Sweep-tone measurement" method is a way of measuring the frequency response characteristics of a system used regularly in acoustic measurements. A sinusoidal function with a specific frequency gives a unique response of the system to that frequency. In the demonstration shown in Figure 3, we constructed a sweep-tone generator using Interactive Physics software. Measuring the formant frequencies of live human speakers, by employing this method, were actually conducted by applying specially designed mechanical vibrator to the speaker's surface of the glottis. [4]

Figure 3 shows the "sweep-tone" measurement of the frequency response of the "uniform" vocal tract. The result shows the characteristics of a neutral vowel. There is "resonance" at frequency with large amplitude in the rows in (b). We call these resonant frequencies, in vowel production, first formant (F1), second (F2), third (F3), and so on, starting from the lowest resonant frequency.

In the result shown in Figure 3, absolute values of masses, spring constants, and frequencies are not explicitly specified and they do not matter. The important point is that, when we discuss resonance (formants), relative relations are maintained.

Assuming that the shape of the "area function" is the same, the ratio of formant frequency values remains unchanged, regardless the sizes of the tract. Therefore, for the uniform tube with one end open and the other end closed, resonant frequencies have ratios of 1, 3, 5, 7, … or odd-number times that of the lowest resonance (F1). Assuming the length of the tract to be 17 cm, they are 500 Hz (F1), 1500 (F2), 2500 (F3), 3500 (F4), etc. as we frequently encounter in the textbook of speech science. In Figures 3, 4, and 5, those values are displayed for reference.

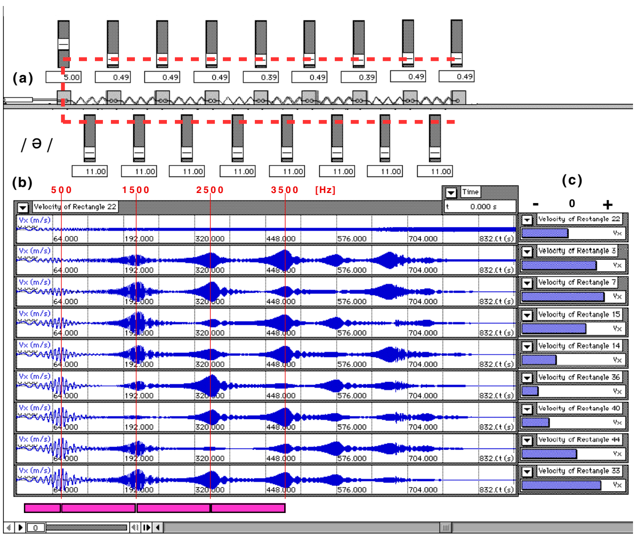

Figure 3 Simulated result for the neutral vowel /ɘ/. A uniform tube is indicated by dotted lines. Rows of waveform display in (b) represent velocities of each mass. Resonance of each mass is represented by the increase in amplitude in velocity during the "sweep-tone" measurement. Bar graphs on the right (in (c)) show a velocity profile for the second formant.

ISBN 1-876346-48-5 © 2003 UAB

3.5 FORMANTS

"Vowel formants", the resonance of a vocal tract configured for vowel production, can also be demonstrated.

Figure 4 shows the result of the simulation of a back vowel /a/. The effect of a narrow constriction close to the glottis is simulated by an increase in both the mass and the spring at the constriction. Notice that the "formant" frequencies, F1 and F2, move closer to each other compared with the "neutral" vowel shown in Figure 3. The magnitude of resonance is shown by wave amplitude of the velocity (in (b)) of masses. Velocity profile corresponding to F2 is shown in (c).

In a similar manner, a front vowel /i/ is simulated in Figure 5. F1 and F2 move apart relative to the "neutral" vowel shown in Figure 3. In this example, it is particularly noticeable that, for the resonance corresponding to F3, the strength of resonance (the magnitude of volume velocity) is distinct in the back cavity compared with the front cavity.

In this kind of simulation, we can obtain information not only about the resonant frequency, but also about the location of dominant resonance. When we talk about the formant frequency, we simply refer to the frequency of "resonance". However, displays such as shown in Figures 3, 4, and 5 also show the patterns of "standing waves", or information about how strong the resonance is in the cavities in front or behind the constriction(s). We can also use the information in discussing vowel articulation.

4. CONCLUSIONS

The physical basics of vowel production have been demonstrated by use of Interactive Physics. By simulating the serially connected mass-spring units, basic notions in acoustics and speech production are easily and effectively demonstrated. This kind of approach is useful not only for educational purposes, but also in research, since we not obtain the information about (formant) frequency but also about the location of resonance whose relation to vowel production is still unknown.

REFERENCES

[1] Interactive Physics, Version 2.5, MSC. Software, 1994. (http://www.interactivephysics.com/home.html)

[2] J. L. Flanagan, *Speech Analysis Synthesis and Perception 2nd Ed.,* New York NY, Springer-Verlag, 1972.

[3] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J.V. Sanders, Fundamantals of Acoustics 3rd Ed., New York NY, John Wiley & Sons, 1982.

[4] O. Fujimura and J. Lindqvist, "Sweeptone Sweep-tone Measurements of Vocal Tract Characteristics," *Journal of the Acoustical Society of America*, vol. 49, pp. 541–558, 1971.

* QuickTime is a registered mark of Apple Computer Inc.

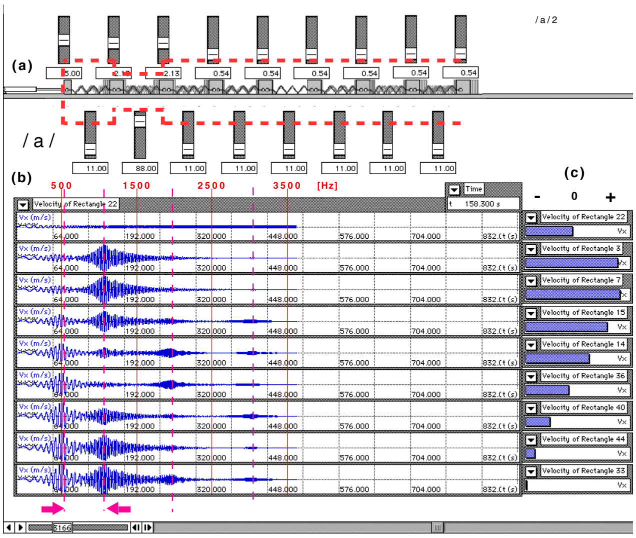

Figure 4 Simulated result of a back vowel /a/. The tube with a constriction, near the glottis (left in (a)), is indicated by dotted lines. Masses and springs at the constriction have greater values simulating the effect of "constriction". Velocity profile corresponding to F2 is shown in (c).

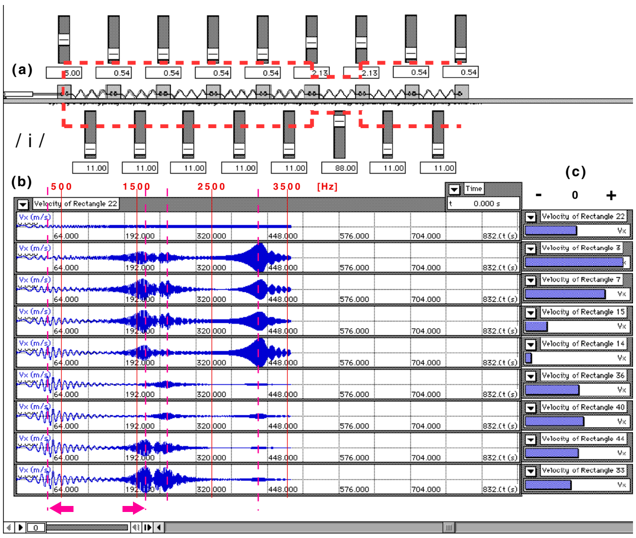

Figure 5 Simulated result for a front vowel /i/. The tube with a constriction, near the lips (right in (a)), is indicated by dotted lines. Masses and springs at the constriction have greater values simulating the effect of "constriction". Velocity profile corresponding to F3 is shown in (c).