Physical and Computer-based Tools for Teaching Phonetics

Takayuki Arai

Dept. of Electrical and Electronics Engineering, Sophia University, Japan

E-mail: arai@sophia.ac.jp

ABSTRACT

Tools for teaching phonetics were considered. We especially focused on two types of tools: physical and computer-based. There are several important aspects of each type. For physical tools, 1) they are real and they enable students to understand the material intuitively; 2) we can show students phenomena in real-time; and 3) students can use them for hands-on experiments. For computer-based tools, 4) we can deal with complex and/or high-speed phenomena virtually; 5) we can visualize unseen phenomena; 6) multi-media environments are possible; and 7) we can effectively use the World Wide Web. In this paper, existing physical and computer-based tools were evaluated in the light of the seven points above. In addition, a new attempt was proposed, an integration of a physical model of the human vocal tract and a computer-based model. Although there are many of tools already in existence, we need to develop tools that are particularly suited for teaching phonetics.

1. INTRODUCTION

These days, a variety of tools for teaching phonetics are widely available including textbooks, recordings, databases, physical tools, and computer-based tools. People are paying more and more attention to education in Acoustics and Speech Science (e.g., [1,2]); we have reported our attempts at several conferences [3-9].

Important aspects of physical and computer-based tools are as follows.

For physical tools,

1) They are real, enabling students to understand the material intuitively;

2) we can show students phenomena in real-time; and

3) students can use them for hands-on experiments.

For computer-based tools,

4) we can deal with complex and/or high-speed phenomena virtually;

5) we can visualize unseen phenomena;

6) multi-media environments are useful (e.g., images / video, audio and texts can be displayed); and

7) we can effectively use the World Wide Web.

In this paper, we will demonstrate each of the above aspects by evaluating existing tools, and we will discuss their features and propose future usage of tools for teaching phonetics.

2. PHYSICAL TOOLS

Physical tools are, in general, useful for intuitive learning, and their demonstration is usually done in real time. Some can be used for hands-on experiments.

Some simple demonstrations of speech aerodynamics are possible with one end of a straw inserted into a glass of water and the other end between the speaker's lips. This demonstrates oral air pressure during various sounds [10].

For demonstrating standing waves, Kundt's experiment is useful. A "Slinky" (coiled spring) is also useful for this purpose (with the concept of traveling and reflection waves). Furthermore, the nodes and antinodes in the standing waves, and the difference between longitudinal and transverse waves, can also be illustrated.

A straw may also be used to demonstrate the Bernoulli effect, that is, the pressure is lower in a moving air stream than in a stationary air stream.

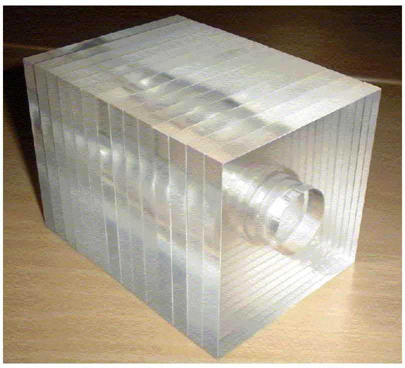

Fig. 1: Physical models of the human vocal tract

 ISBN 1-876346-48-5

Showing measurements related to phonation and articulation may also be instructive. The PCquirer & Macquirer systems allow one to measure aerodynamic data (e.g., air pressure and oral / nasal airflow [11]. In addition, electroglottograph (EGG) may be used to evaluate fundamental frequency and voice quality, and an electropalatograph for place and manner of articulation [12,13].

The examples so far are intuitive and are performed in real time. The use of real-time teaching tools is valuable because students can directly observe the association between a cause and resulting sounds. Spectrograms can also be obtained in real time (e.g., Kay DSP Sona-Graph [13]).

Other examples of real-time demonstrations which are extremely important for intuitive learning are physical models of the human vocal tract. Figure 1 shows some examples of the models we have developed recently. Designing such models is not new (see [14]); Chiba and Kajiyama (1941) made models based on their measurement [15]. Umeda and Teranishi (1966) also made a physical model with several sliding bars, so that an arbitrary configuration of the vocal tract can be implemented [16]. Some museums have physical models of the human vocal tract (e.g., Exploratorium in San Francisco under supervision of Prof. John Ohala [17]).

In 2001, Arai replicated Chiba and Kajiyama's precise models and showed that they are extremely effective in the class [3]. Arai (2001) proposed a simple but powerful demonstration of vowel production using physical models ([3-9]; see also [18]) and an electrolarynx, an effective tool for demonstrating the source filter theory [19].

3. COMPUTER-BASED TOOLS

3.1 Software tools for analysis/synthesis

Many speech tools analyze speech sounds and display results, such as FFT / LPC spectrum, spectrogram, formant tracking, f0/power contour. Some of them can make formant plots (PlotFormants [20]).

Modifying speech sounds allows one to observe which speech parameters govern speech perception. Using the Computerized Speech Lab. (CSL) with ASL for LPC resynthesis, one may change various segmental and supra-segmental parameters [13]. Recently, Praat is widely used for education [21].

The Klatt synthesizer may also be used to demonstrate the importance of formants [22]. With the HLsyn, a high-level formant sythesizer, one may manipulate articulatory parameters [23]. Applying a Text-to-Speech synthesizer, such as Mbrola [24], is educational, especially for teaching the role of prosody.

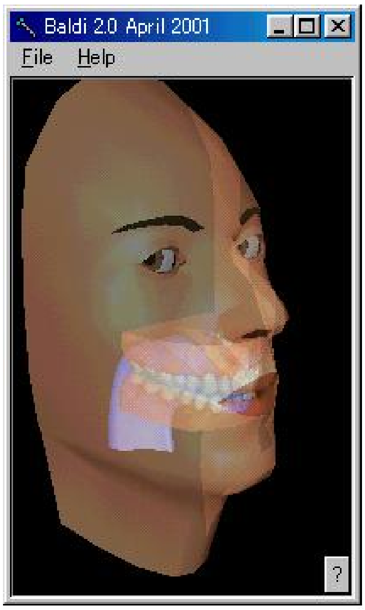

Fig. 2: Baldi in the CSLU Tookit.

3.2 Quasi-real-time tools

As mentioned in Section 2, real-time demonstration is key. Some software displays on-line spectrograms in real time (SpeechStation2 [23]).

3.3 Multi-media environments

In multi-media environments, several modalities can be used at the same time, such as images, video, audio and texts. The Sounds of the World's Languages illustrates the sounds of over 100 languages under investigation at UCLA [20,25,26]. Many sounds of various languages by native speakers are available, annotated with the symbols of the IPA (International Phonetic Alphabet), so that students can practice to pronunciation and auditory discrimination.

Speech Production and Perception I is a CD-ROM course in speech science designed for undergraduates [23]. Phonetic & Perception Simulation Programs [13] is an educational tool for acoustic phonetic and communication disorders sciences. Since these multimedia tools are interactive, students can learn materials at their own paces. The Mouton Interactive Introduction to Phonetics and Phonology helps students to learn about phonetics and phonology [27].

To explain the interaction between audio and visual information of speech, the McGurk effect may be the best example [28,29].

Phonetic Sammy is a program on the Web for specifying articulations and seeing the resulting vocal tract shape and the corresponding IPA symbol [30].

3.4 Corpus and Labeling

Collecting speech data and compiling speech corpora / databases (such as, [31-33]), are meaningful not only for research purposes but also for education. Labeling phonetic / phonemic information as well as prosodic [34] and other linguistic information pertaining to the speech data is also crucial.

3.5 Visualization

With computer simulation, we can deal with complex and/or high-speed phenomena by applying a visualization technique. For example, revealing speech organs on a computer display

ISBN 1-876346-48-5

which are normally hidden from view may be educational for students, in addition to the "Anatomical Charts & Models."

A talking head such as "Baldi" in the CSLU Toolkit [32] allows us to see the movement of articulators through the skin (Fig. 2). By changing a parameter, the skin becomes transparent, so that users can determine where the tongue is located, etc. This feature is also useful for second language learners and hearing-impaired children [35].

3.6 Utilizing the World Wide Web

The OpenCourseWare project at MIT provides the Web-based course "Laboratory on the Physiology, Acoustics, and Perception of Speech" for all to use including cineradiographic motion pictures [36]. There are more movies of the vocal tract and vocal cords in action in [37-39].

4. INTEGRATING PHYSICAL AND COMPUTER-BASED TOOLS

As shown in Sections 2 and 3, both physical and computer-based tools are advantageous for teaching phonetics. If we integrate them, great benefit may be expected. Figure 3 shows an example of such integration. On a computer, we can simulate a vowel sound from the area function of a human vocal tract [40]. From the same data, we can also design a physical model of the vocal tract, which also produces a vowel sound. By comparing the two signals, students are able to determine whether a physical phenomenon is supported by a theory.

5. SUMMARY

Existing physical and computer-based tools were evaluated. In addition, something new was proposed, an integration of a physical and a computer-based tool. Although there are many of tools already in existence, we need to continue to develop tools and education methods that are particularly suited for teaching phonetics.

ACKNOWLEDGMENT

I would like to thank all of the people who helped me in various ways, especially Dawn Behne, Robert Berkovitz, Suzanne Boyce, Pam Connors, Helen Hanson, Michel Jackson, Tomoko Kitamura, Terri Lander, Shinji Maeda, John Ohala, Ralph Ohde, Daniel Recasens, Takahito Shinya, Janet Slifka, Chris Soland, Maria-Josep Sole, Ken Stevens, Jacqueline Vaissiere, and Natasha Warner.

REFERENCES

[1] http://www.dcs.shef.ac.uk/spandh/ed_arena/

[2] http://www.ling.umu.se/arena/

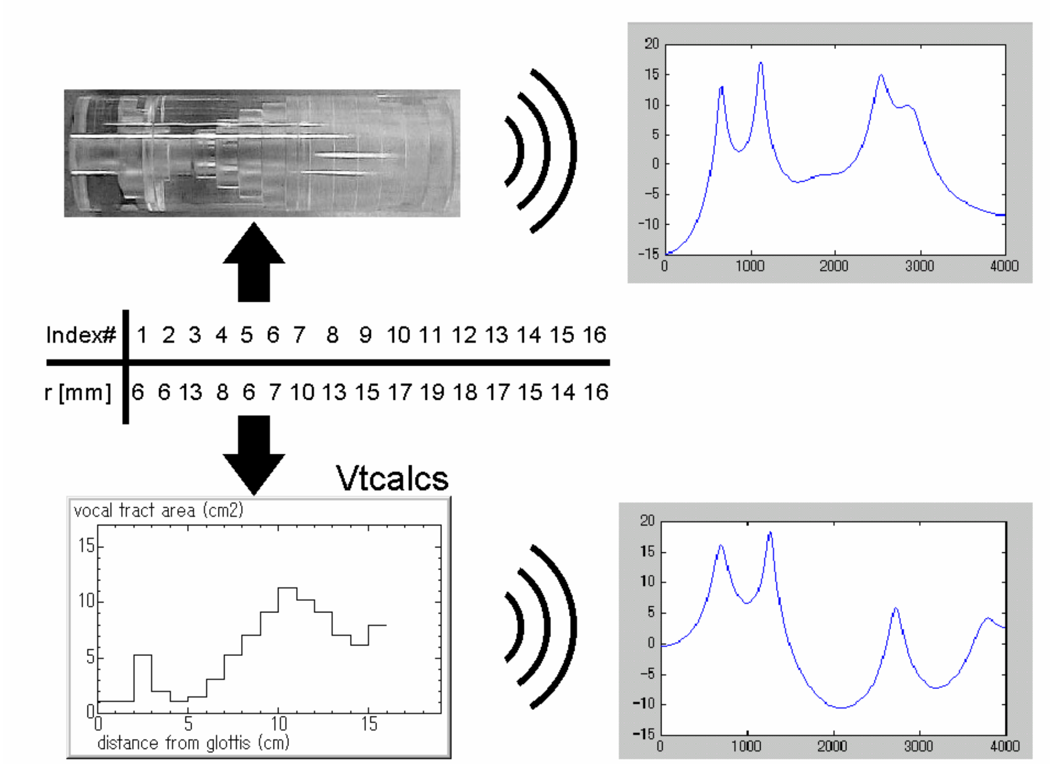

Fig. 3: Integrating physical and computer-based tools.

ISBN 1-876346-48-5 © 2003 UAB

[3] T. Arai (2001) "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," *Journal of the Phonetic Society of Japan*, 5(2):31-38.

[4] T. Arai, N. Usuki and Y. Murahara (2001) "Prototype of a vocal-tract model for vowel production designed for education in speech science," *Proc. of Eurospeech*, 4:2791-2794, Aalborg.

[5] T. Arai (2002) "An effective method for education in acoustics and speech science: Integrating textbooks, computer simulation and physical models," *Proc. of the Forum Acusticum Sevilla*.

[6] N. Saika, E. Maeda, N. Usuki, T. Arai and Y. Murahara (2002) "Developing mechanical models of the human vocal tract for education in speech science," *Proc. of the Forum Acusticum Sevilla*.

[7] E. Maeda, N. Usuki, T. Arai and Y. Murahara (2002) "The importance of physical models of the human vocal tract for education in acoustics in the digital era," *Proc. of China-Japan Joint Conference on Acoustics*, 163-166, Nanjing.

[8] T. Arai, E. Maeda, N. Saika and Y. Murahara (2002), "Physical models of the human vocal tract as tools for education in acoustics," *Proc. of the First Pan-American/Iberian Meeting on Acoustics*, Cancun.

[9] E. Maeda, T. Arai, N. Saika and Y. Murahara (2002) "Lab experiment using physical models of the human vocal tract for high-school students," *Proc. of the First Pan-American/Iberian Meeting on Acoustics*, Cancun.

[10] R. D. Kent and C. Read (1992) *The Acoustic Analysis of Speech*, Singular Pub., San Diego.

[11] http://www.sciconrd.com/xmulti.htm

[12] http://www.laryngograph.com/

[13] http://www.kayelemetrics.com/

[14] B. Gold and N. Morgan (2000) *Speech and Audio Signal Processing*, John Wiley & Sons.

[15] T. Chiba and M. Kajiyama (1941) *The Vowel: Its Nature and Structure*, Tokyo-Kaiseikan Pub. Co., Ltd., Tokyo.

[16] N. Umeda and R. Teranishi (1966) "Phonemic feature and vocal feature: Synthesis of speech sounds, using an acoustic model of vocal tract," *J. Acoust. Soc. Jpn.*, 22(4):195-203.

[17] http://www.exploratorium.edu/exhibits/vocal_vowels/

[18] T. Lander and T. Arai (2003) "Using Arai's vocal tract models for education in Phonetics," *Proc. of ICPhS*, Barcelona.

[19] G. Fant (1960) *Acoustic Theory of Speech Production*, Mouton, The Hague.

[20] http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/acoustic.html

[21] http://www.fon.hum.uva.nl/praat/

[22] D. H. Klatt (1980) "Software for a cascade / parallel formant synthesizer," *J. Acoust. Soc. Am.*, 67:979-995.

[23] http://www.sens.com/

[24] J. Vaissiere (2003) "Some useful tools for teaching the segmental and suprasegmental aspects of speech," *Proc. of ICPhS*, Barcelona.

[25] P. Ladefoged (2000) *Vowels and Consonants: An Introduction to the Sounds of Languages*, Blackwell Pub., Oxford.

[26] P. Ladefoged (2001) *A Course in Phonetics*, Heinle & Heinle Pub., Boston.

[27] http://www.uni-marburg.de/linguistik/dgweb/phonology/

[28] http://www.media.uio.no/personer/arntm/McGurk_english.html

[29] http://www.psych.ucr.edu/avspeech/lab

[30] http://www.chass.utoronto.ca/~danhall/phonetics/sammy.html

[31] http://www.ldc.upenn.edu/

[32] http://cslu.cse.ogi.edu/toolkit/

[33] http://www.speechtech.com/main.html

[34] http://www.ling.ohio-state.edu/~tobi/

[35] P. Connors, A. Davis, G. Fortier, K. Gilley, B. Rundle, C. Soland and A. Tarachow (1999) "Participatory design: Classroom applications and experiences," *Proc. of the ICPhS*, 1285-1288, San Francisco.

[36] J. Slifka, S. Shattuck-Hufnagel and L. Koller (2003) "Speech on the Web: An MIT lab course," *Proc. of ICPhS*, Barcelona.

[37] http://www.phon.ox.ac.uk/~jcoleman/phonation.htm

[38] http://www.humnet.ucla.edu/humnet/linguistics/faciliti/demos/vocalfolds/vocalfolds.htm

[39] http://www.unc.edu/~jlsmith/pht-url.html

[40] S. Maeda (1982) "A digital simulation method of the vocal-tract system," *Speech Communication*, 1:199-229.