

# Implementation and Usability Study of a Multimedia Phonetic Teaching and Recognition System

Minder Chen<sup>1\*</sup>, Yun-Long Lay<sup>2</sup>, Hui-Jen Yang<sup>3</sup>

## ABSTRACT

Learning Mandarin phonemes are difficult for hearing impaired, children, and non-native speakers. A multimedia phonetic learning system has been implemented to help aforementioned people. The system provides an interactive multimedia learning environment to help learners practice the phonetic pronunciation more lively. This system can show multimedia content that includes a video demo, sequence of mouth shapes, and place of articulation diagram. The system is equipped with a microphone to record the learner's phonetic pronunciation. The Hidden Markov Model is used to recognize each pronunciation. This recognition module provides proper feedback to learners to improve their learning performances. An empirical study has been conducted to test the usability of the system. The experiment is designed to evaluate the intention of system use based on technology acceptance model (TAM) and innovation diffusion theory (IDT). The results of the empirical study are presented. The implications of the empirical study are discussed in the conclusion of the paper.

## I. INTRODUCTION

A previous study found that information can be better retained if it is presented through visual images, especially via animated images (Paivio, 1971). A multimedia system provides a rich language learning environment containing texts, sounds, pictures/diagrams, and videos. Previous research demonstrated that multimedia-based learning tools can support and facilitate self-directed learning (Mayer, 2005). The research of Yang et al. (2005) showed that learning phonetics is usually based on the learner's observation of the shapes of teacher's mouth while pronouncing words. The student practiced repeatedly by imitating the teacher's pronunciations. However, abstract phonetic symbols are difficult for students to memorize. The learning process is boring and inefficient. Therefore, in this research, we have implemented a computer-aided multimedia learning system for learning phonemes. The system provides instructional materials consisting of videos,

images, sounds, and texts. It also integrates a voice recognition module to provide feedback to the learner. Harless et al. (1999) pointed out that using multimedia and voice recognition technology can help people learn pronunciation more effectively. Through the integration of various engaging media formats in the system, we expect that it may increase the learner's motivation and the quality of instructions. Mayer (2001) argued that using both text and pictures can help people understand the language instructional materials better. Specifically, when text and pictures are used together, learners can make better connections between texts and pictures. Therefore, it would increase learners' interest and motivation to learn the subject. Mayer et al. (2003) found that the learners have better learning outcomes when texts and pictures were used together. Actually, with progress in information technology, computer-aided learning systems have been successfully developed (Youdelman & Levitt, 1991; Yang et al., 2005). In recent years, many of the computer-aided learning systems have added computer games or animations to attract the learner's attention (Youdelman and Levitt, 1991; Yang et al., 2005). In this paper, we present the implementation of a system called Multimedia Phonetic Teaching and Recognition System (MPTRS). The system is designed to help hearing impaired, children, and foreigners learn Chinese Phonetic Symbols. The design and implementation of the system is first presented. An empirical study has been conducted to test the usability of the system. The experiment evaluated the intention of system use based on technology acceptance model (TAM) and innovation diffusion theory (IDT). The results of the empirical study are presented.

## II. THEORETICAL BASES

The MPTRS we implemented has a recognition subsystem that can provide feedback to the learner. The Hidden Markov Model is used to implement the speech recognition function and it is explained in Section II.1. The empirical study of the system based on Technology Acceptance Model and Diffusion of Innovation Theory are explained in Section II.2.

### 1. Hidden Markov Model

A Hidden Markov model (HMM) is assumed to be a Markov process with unobserved (hidden) states. There are several well-known algorithms for implementing HMM applications. For instance, the Viterbi algorithm is used to

Paper submitted 08/15/11; revised 11/15/11; accepted 12/30/11.

\* Corresponding Author: Minder Chen (E-mail: minder.chen@csuci.edu).

<sup>1</sup> Martin V. Smith School of Business and Economics, California State University Channel Islands, USA

<sup>2</sup> Department of Electronic, National Chin-Yi University of Technology, Taiwan

<sup>3</sup> Department of MIS, National Chin-Yi University of Technology, Taiwan

compute the most-likely corresponding sequence of states; the forward algorithm is used to compute the probability of the observation's sequence; the Baum-Welch algorithm is used to estimate the starting probabilities, the transition function, and the observation function (Ryan and Nudd, 1993; Wang et al., 2007). The Markov process is a time-varying random phenomenon for which a specific Markov property holds (Cappe et al., 2005). An HMM can be considered as the simplest dynamic Bayesian network. The regular Markov model assumes the Markov property which models the state of a system with a random variable that changes over time. An HMM is a Markov chain with unobserved (hidden) states. The visible output of the model depends on these states. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM reveals some information about the sequence of states (Cappe, Moulines, and Ryden, 2005). HMM has been used in many applications such as voice recognition, gesture recognition, and part-of-speech tagging. HMM is used to implement the recognition function in MPTRS.

## 2. Technology Acceptance Model and Diffusion of Innovation Theory

The technology acceptance model (TAM) and diffusion of innovation theory (DIT) are two well-established theories that have been used successfully to explain the adoption of various technologies and products (Davis, 1989; Roger, 2003; Chen & Tan, 2004). Previous research suggested that TAM integrated with innovation diffusion theory can improve its predictive power of system use (Chen & Tan, 2004). Two salient constructs, perceived ease of use (PEOU) and perceived usefulness (PU), in the TAM model are the primary antecedents to predict the intention of system use. Additional factors have been added to the original TAM, such as anxiety or playfulness, to examine the use intention of newly developed systems (Yang et al., 2005; Yang et al., 2006; Yang et al., 2008).

Another popular theory explaining user adoption is IDT (Roger, 2003). IDT believes that innovation diffusion is achieved through user's acceptance or the use of new things or new systems (Chen & Tan, 2004; Yang et al., 2006). IDT identified five categories of adopters for new systems or technologies: innovators, early adopters, early majority, late majority, and laggards. It also identified five innovation adoption factors: relative advantage, compatibility, complexity, triability, and observability. Tomatzky et al. (1982) found that relative advantage, compatibility, and complexity are the important factors that affect the rate of innovation adoption.

## III. METHODS

### 1. Pronunciation Categories

Chinese (Mandarin) phonetic symbols consist of

consonants, vowels, and tones. In Chinese, each character corresponds to one syllable. A consonant known as initial often appears in the front of a syllable. The remaining one or two phonetic symbols are the final vowel(s). The phonetic pronunciation learning lessons implemented in MPTRS covers only consonants. These consonants are categorized into seven types based on their "pronunciation position" and "pronunciation method" as shown in Table 1 (Lo, 2008). A consonant is a speech sound that is articulated with complete or partial closure of the vocal tract and is pronounced with the air flowing out from the lungs through the glottis. Pronunciation position (i.e., place of articulation) refers to the part of the mouth where the airstream is obstructed when a consonant is pronounced (Wikipedia, 2012). For instance, the pronunciations of ㄅ(Bo), ㄆ(Po), and ㄇ(Mo) use bilabial to block the air flow while ㄉ(De), ㄊ(Te), ㄋ(Ne), and ㄌ(Le) use the alveolar and tongue to stop the air flow. Articulation is the physical production of a particular speech sound. It refers to the air-flow through the vocal organs when a sound is pronounced. The tongue position and mouth shape affect the vocal cord vibrations to produce different sounds as shown in Figure 1.

Table 1. Chinese Phonetic Symbols (Consonants)

Types	Symbols
1. Bilabial	ㄅ(Bo)、ㄆ(Po)、ㄇ(Mo)
2. Labio-dental	ㄈ(Fo)
3. Alveolar	ㄉ(De)、ㄊ(Te)、ㄋ(Ne)、ㄌ(Le)
4. Velar	ㄍ(Ge)、ㄎ(Ke)、ㄏ(He)
5. Palatal	ㄐ(Ji)、ㄑ(Qi)、ㄒ(Xi)
6. Retroflex	ㄗ(Jr)、ㄘ(Chr)、ㄙ(Shr)、ㄨ(Jih)
7. Lingua-dental	ㄗ(Tz)、ㄘ(Tsz)、ㄨ(Sz)

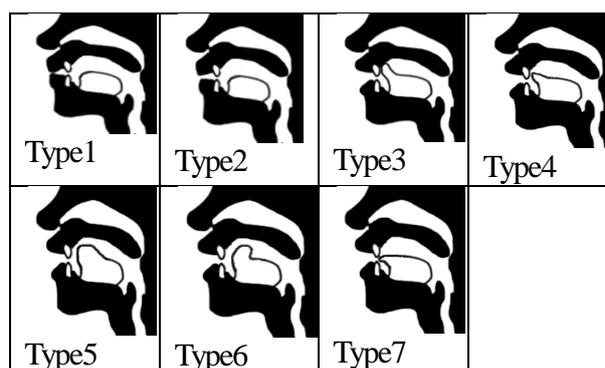


Figure 1. Seven types of consonant pronunciation's tongue positions and mouth shapes

### 2. Phonetic Recognition

The phonetic recognition function relies on the features extracted from the sound pronounced (Wang, 2009). The voice stored in the database will be processed to extract suitable characteristic parameters as a basis for identification. First, we find out the starting point and ending point of the voice through the energy detection. Then for the identified voice signal, the system takes 240

sample points in a frame and then conducts the following analyses sequentially: Pre-emphasis, Hamming window, Autocorrelation coefficients analysis, Linear Predict Coding (LPC) analysis, and Cepstrum coefficients. Each frame gets ten Cepstrum coefficients as a group of feature coefficients vector  $x_j$ . Put a pronunciation for all frames sequentially divided into three groups ( $j_1, j_2, j_3$ ) as three initial state ( $S_1, S_2, S_3$ ). Table 2 is one of groups to express the state which has  $j$  frame and each frame has  $i$  ( $i=10$ ) Cepstrum coefficients represented by  $C_{ij}$ . We apply the Hidden Markov Model (Rabiner & Juang, 1986) to calculate the state expectation matrix of each group  $\mu_i$  (Eq. 1) and the Variance matrix  $R_i$  (Eq. 2). After that, the Gaussian mixture model (Ramalingam & Krishnan, 2006) (Eq. 3) is passed the state  $\mu_i$  and  $R_i$  respectively to get the corresponding initial state probability function value in each frame, shown in Table 3. The probability function of the initial value applying Viterbi algorithm (Viterbi, 2006; Lou, 1995) is used to find out the optimal state sequence. The main purpose of Viterbi algorithm is to obtain a set of input speech which is most similar to the model portfolio as the recognition model.

$$\mu_i = (C_{i1} + C_{i2} + C_{i3} + \dots + C_{ij}) / j \tag{1}$$

$$R_i = (C_{i1}^2 + C_{i2}^2 + C_{i3}^2 + \dots + C_{ij}^2) / j - \mu_i^2 \tag{2}$$

$$G(x_j) = (2\pi)^{-\frac{N}{2}} |R_i|^{-\frac{1}{2}} \exp[-\frac{1}{2}(x_j - \mu_i)^T R_i^{-1} (x_j - \mu_i)] \tag{3}$$

**Table 2. One of the States within the Frame and Cepstrum coefficients**

Frame	1	2	3	4	...	j
Cepstrum coefficients	$C_{11}$	$C_{12}$	$C_{13}$	$C_{14}$		$C_{1j}$
	$C_{21}$	$C_{22}$	$C_{23}$	$C_{24}$		$C_{2j}$
	$C_{31}$	$C_{32}$	$C_{33}$	$C_{34}$		$C_{3j}$
	:	:	:	:		:
	$C_{i1}$	$C_{i2}$	$C_{i3}$	$C_{i4}$	$C_{ij}$	

**Table 3. All Frame probability function of the initial value corresponding to State**

Frame	1	2	3	...	J
State <sub>1</sub> ( $j_1$ )	$G(x_1)$	$G(x_2)$	$G(x_3)$		$G(x_j)$
State <sub>2</sub> ( $j_2$ )	$G(x_1)$	$G(x_2)$	$G(x_3)$		$G(x_j)$
State <sub>3</sub> ( $j_3$ )	$G(x_1)$	$G(x_2)$	$G(x_3)$		$G(x_j)$

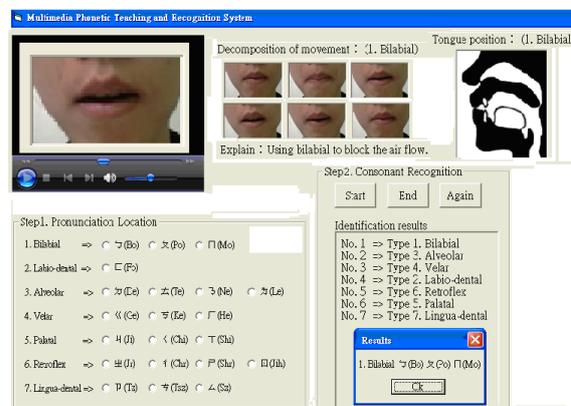
### IV. SYSTEM IMPLEMENTATION

The system architecture has three key elements: teaching, practice, and voice recognition. We developed teaching materials for the precise pronunciation of each consonant consisting of lips diagram, text description, and tongue position image. Pictures and recorded sound have been converted to multimedia file formats.

The typical usage scenario of MPTRS is described in the following. The user first selects the learning phoneme. The system shows the demonstration video and a series of mouth moving pictures to explain the correct pronunciation. The user could articulate the phonemes for the pronunciation practice. The last stage is to evaluate the learning performance of the phonetic recognition system.

In the following scenario, the  $\text{ㄗ}$  (Po) consonant's sound is used as an example to explain the initial phonetic word training. An effective voice sample rate is at least 8 kHz or more (Tabbers et al., 2004). Therefore, the system uses 8 kHz and 16 bit of sample rate to record the learner's voices. When the users pronounce the  $\text{ㄗ}$ (Po) sound, the length is about 1 second. It will generate 104 frames via the aforementioned process. After removing the sound header, 102 frames have been generated. The system applies the HMM method to calculate three states as the recognition features. Each state has 34 frames to individually calculated  $\mu_i$  (Eq. 1) and  $R_i$  (Eq. 2) and then substituted them into Eq. 3 to set up each frame's corresponding initial value of the frame's probability function. To calculate the maximum possible weighted values of the path, this value is the combination of the three state features and it is used to compare with the training sample using the Viterbi algorithm.

This system works with all versions of Microsoft Windows Systems. The basic devices include a personal computer with proper audio Input/Output devices (i.e., a microphone, a speaker). Chinese phonemes have 21 consonant sounds which can be classified into seven categories of articulation. The user interface of the Multimedia Phonetic Teaching and Recognition System (MPTRS) is shown in Figure 2. The typical scenario of a user using the MPTRS is discussed in the following:



**Figure 2. Multimedia Phonetic Teaching and Recognition System**

- (1) In accordance with the articulation category, the user first clicks on a mandarin phonetic symbol. Then the user can read pronunciation explanations, study decomposition of tongue movement and location map, as well as view multimedia demonstration teaching.
- (2) The user can press the recording button to record his or her pronunciation.
- (3) In the recognition area, the user can press the “start” button to identify whether the pronunciation is suitable or not, or return to step 1 again.

The user could use the system over and over again to watch the teaching of different phonetic articulations. They can also use the system’s phonetic recognition function to receive feedbacks. A user’s pronunciation of a consonant (i.e., a testing sample) could be captured and compared to the sample data stored in the system. After calculating the weighted value of a testing sample and comparing it with the sample data, the best possible matched pronunciation category could be found. The phonetic category identification rate is close to 85%. The phonetic recognition function offers the learner a self-assessment tool. The consonant recognition accuracy rate of the system is shown in Table 4.

**Table 4. Consonant recognition accuracy rate**

Types		Accuracy rate
Bilabial	ㄅ(Bo)、ㄆ(Po)、ㄇ(Mo)	80%
Labio-dental	ㄈ(Fo)	85%
Alveolar	ㄉ(De)、ㄊ(Te)、ㄋ(Ne)、ㄌ(Le)	80%
Velar	ㄍ(Ge)、ㄎ(Ke)、ㄏ(He)	80%
Palatal	ㄐ(Ji)、ㄑ(Chi)、ㄒ(Shi)	90%
Retroflex	ㄗ(Jr)、ㄘ(Chr)、ㄙ(Shr)、ㄨ(Jih)	100%
Lingua-dental	ㄗ(Tz)、ㄘ(Tsz)、ㄨ(Sz)	95%

## V. EVALUATION EXPERIMENT

### 1. Hypotheses

We conducted an empirical study to evaluate the usefulness and usability of MPTRS. Two widely accepted theories in technology adoption, TAM and IDT, are used in the empirical study. IDT identifies five characteristics that can be used to determine the success of innovation diffusion. These five characteristics are relative advantage, complexity, compatibility, observability, and trialability (Roger, 1993; Robinson, 2009). Relative advantage (RA) refers to a new innovation surpassing current practices. Complexity refers to the perceived difficulty of learning to use and understand a new system or technology. Compatibility means an innovation is perceived as consistent with the adopter’s existing values, past experiences, and needs. Observability refers to the results of an innovation being visible and understood by others. Trialability refers to an innovation that may be experimented with on a limited basis. The constructs in

IDT overlaps with TAM’s constructs to a certain degree. RA construct in IDT is often treated as the same construct of perceived usefulness in TAM. Complexity construct in IDT is often treated as the reverse construct of ease of use. Therefore, in this study RA and complexity is used as surrogates of perceived usefulness and ease of use, respectively. Previous studies have found that complexity and RA are important factors in technology adoption decisions (Tornatzky et al., 1982; Chen & Tan, 2004). Additional studies have also found that playfulness and anxiety have a significant impact on the behavioral intention to use the learning system (Yang et al., 2007; Yang et al., 2008). Therefore, in this study complexity and RA are defined as “the degree of recognition learning system is perceived to be difficult to operate” and “the recognition learning system is superseding the current mouse system”, respectively. Playfulness is defined as “the degree of recognition learning system is perceived to be interested”. System anxiety is defined as “the degree of recognition learning system is perceived to be afraid”. Dependence intention is defined as “the degree of system is perceived to rely on”. The scales of RA, complexity, playfulness, system anxiety and dependence intention are all modified from previous research (Davis, 1989; Chen, 1999; Yang et al., 2005; Yang et al., 2006). Five items for RA and complexity respectively, 7 items for system anxiety, three items for playfulness, and 7 items for dependence intention are used in this study’s questionnaire.

From the aforementioned well-established theoretical basis, the following hypotheses were proposed to test the system’s usability:

H1: Relative advantage of system has a positive influence on dependence intention of system.

H2: Complexity of system has a negative influence on dependence intention of system.

H3: System anxiety has a negative influence on dependence intention of system.

H4: Playfulness of system has a positive effect on dependence intention of system;

### 2. Data Collection

A questionnaire survey was conducted to test the dependent intention of system. Teachers from National Taichung Deaf School were recruited to administrate the empirical study. The experiment was conducted by following the steps: (1) the system was installed at the school; (2) teachers showed their students who are first and second grade in the elementary school how to operate the system; (3) one month later, students were asked to answer the questionnaire using a five-point Likert scale. 36 valid questionnaires were complete and usable for data analysis.

The reason for including only the first grade students in the evaluation stage is that students in the first grade are required to learn Mandarin phonemes. Learning Mandarin phonemes is the basic first step for students to learn the pronunciation of Chinese words. The survey was conducted after students had at least one-month experience with using MPTRS.

## VI. RESULTS

Table 5 shows the regression results of four hypotheses. The Beta, t value and significant of each variable are indicated in the Table 5. All four hypotheses are supported in the system usability study. The results found that these factors separately have significant influence on the intention to use MPTRS. This study proves that system complexity has a negative impact on dependent intention (Beta= .477,  $p<0.001$ ). This study also confirms that relative advantage has a positive influence on dependent intention (Beta= .455,  $p<0.01$ ). Additionally, this study also demonstrates that system playfulness has a positive effect on dependent intention (Beta=.450,  $p<0.05$ ). Furthermore, this study also presents that system anxiety has a negative impact on dependent intention (Beta= -.520,  $p<0.001$ ). Among them, system anxiety has most significant effect on dependent intention, system complexity is the next, followed by relative advantage, and playfulness is the last. This result indicated that hearing impaired students were not afraid of the system and found the system was easy to operate. They also perceived the function of system was better than the existing system and the interface of the system is interesting. They demonstrated high intention to use the system to learning their phonemes.

**Table 5: The Regression of the Hypotheses**

Variables	Beta	t	Sig.
Complexity	-.477	-3.485	.000**
Relative Advantage	.455	2.979	.006**
Playful	.450	2.577	.015*
System anxiety	-.520	-3.663	.000***

Dependent variable: Dependence Intention of the system;  
df=4, F=20.189; Sig=.000; R=.754; R2=.688  
\*\*\*P<0.001, \*\*P<0.01, \*P<0.05

## VII. CONCLUSIONS

This study used a multimedia-based system and HMM method to design a phonetic learning system to attract the user's interest and attention. This system shows the decomposed lip action and moving tongue location for each phonetic symbol. All the images with sounds were merged to be an AVI file and used by the MPTRS. In practice, when MPTRS is used for phonetic teaching, learners can study the lively multimedia teaching materials repeatedly. The phonetic recognition function in MPTRS can provide useful feedback to learners to improve their learning performance.

The 21 consonants of Mandarin phonemes are classified into seven groups. A fixed three states of HMM used has achieved good recognition results despite the fact that pronunciation is easily affected by external factors (e.g., noise in the background environment) or internal factors (e.g., the learner has a cold). We are working on a more advanced phonetic recognition model to reach higher

recognition rate. We are also exploring technologies that will make MPTRS Web-enabled to reach wide audience.

A self-report questionnaire was used to measure the students' intention to use the MPTRS. The result revealed that subjects found that the MPTRS is easy to use, interesting, and useful. They were not afraid of the MPTRS; hence there is high intention to use the system. Based on the results of this empirical study, we conclude that when the system designers develop a computer-aided instruction system, they should make the system simple to use and playful, incorporate multimedia content, and provide appropriate feedback with performance measurement to enhance users' motivation to use the newly implementing system. A limitation in our usability study is that only 36 hearing-impaired subjects participated in the study. The results may not be generalized to other types of users. We are planning to include other types of users in our future studies to make the evaluation more comprehensive.

## REFERENCES

- [1] Cappe, O., Moulines, E., and Ryden, T. (2005). Inference in Hidden Markov Models. Springer.
- [2] Chen, L. D. and Tan, J. (2004) Technology adaptation in E-commerce: key determinants of virtual stores acceptance. *European Management Journal*, 22(1), 76-86.
- [3] Chen L. D. and Tan J. (2004). Technology adaptation in E-commerce: key determinants of virtual stores acceptance. *European Management Journal*, 22(1), 76-86.
- [4] Chen, K. C. (1999). Factors that motivate Internet users to use business-to-customer electronic commerce. Doctoral Dissertation, Cleveland State University Cleveland, OH, USA
- [5] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- [6] Harless, W. G., Zier, M. A., and Duncan, R. C. (1999). Virtual dialogues with native speakers: the evaluation of an interactive multimedia method. *CALICO Journal*, 16(3), 313-337.
- [7] Lo, J. C. (2008). Phonetic symbols teaching (In Chinese). Taipei, Taiwan: China Overseas Chinese Affairs Commission Correspondence School Publishing.
- [8] Lin, C. S. (2001). Digital signal-image and voice processing. Taipei, Taiwan: Chan-Hwa Technology Book Co. (In Chinese).
- [9] Lou, H. L. (1995). Implementing the Viterbi algorithm. *IEEE Signal Processing Magazine*, 12(5), 42-52.
- [10] Mayer, R. E. (2001). *Multimedia learning* Second Edition. New York, NY: Cambridge University Press.
- [11] Mayer, R. E. and Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychology*, 38(1), 43-52.

- [12] Mayer, R. (2005). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.
- [13] Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: Cambridge University Press.
- [14] Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
- [15] Reed, S. K. (2006). Cognitive architectures for multimedia learning. *Educational Psychology*, 41(2), 87-98.
- [16] Ramalingam, A. and Krishnan, S. (2006). Gaussian mixture modeling of short-time Fourier transform features for audio fingerprinting. *IEEE Transactions on Information Forensics and Security*, 1(4), 457-463.
- [17] Rabiner, L. and Juang, B. (1986). An introduction to the Hidden Markov Model. *IEEE ASSP Magazine*, 3(1), 4-16.
- [18] Rogers, E. M. (2003). *Diffusion of innovation*, 5th edition, New-York: The Free Press.
- [19] Ryan, M.S. and Nudd, G.R. (1993). The Viterbi algorithm. Technical report RR-238, University of Warwick, UK.
- [20] Tabbers, H. K., Martens, R. L., and van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cuing. *British Journal of Educational Psychology*, 74, 71-81.
- [21] Tomatzky L. G. and Klein K. J. (1982). Innovation characteristics and innovation adoption-implementation: a meta-analysis of the empirical findings. *IEEE Transactions on Engineering Management* 29(1), 28-45.
- [22] Viterbi, A. J. (2006). A personal history of the Viterbi algorithm. *IEEE Signal Processing Magazine*, 23(4), 120-142.
- [23] Weiss, I., Kramarski, B., and Talis, S. (2006). Effects of multimedia environments on kindergarten children's mathematical achievements and style of learning. *Educational Media International*, 43(1), 3-17.
- [24] Wang, S. C. (2009). *Voice signal handling*. Taipei: All-hua Book Co. (In Chinese).
- [25] Youdelman, K. and Levitt, H. (1991). Speech training of deaf students using a palatographic display, *Proc. of First Intern. Symp. on Speech and Hearing Sciences*, 1-11, Osaka, Japan.
- [26] Yang, H. J. and Lay, Y. L. (2005). Implementation and evaluation of computer-aided Mandarin phonemes training system for hearing-impaired students. *British Journal of Educational Technology*, 36(3), 537-551.
- [27] Yang H. J., Lay, Y. L., and Tsao, C. H. (2006). An Implementation and usability evaluation of automatic cash-payment for hospital, *Journal of Scientific & Industrial Research*, 65, June, 485-494.
- [28] Yang, H. J., Tsao, W. Y., Lay, Y. L., Chen, M., and Liou, Y. C. (2008) Prior Language Experience and Language Anxiety as Predictors for Non-Native Language Commercial Web Site Use Intention, *International Journal of Human-Computer Studies*, 66, Sept., 678-687.
- [29] Wang, Y. W., Giruka, V. C., and Singhal, M. (2007). A Truthful Geographic Forwarding Algorithm for Ad-hoc Networks with Selfish Nodes, *International Journal of Network Security*, 5(3), 252-263.
- [30] Wikipedia (2011), Place of articulation, Dec. 30, [http://en.wikipedia.org/wiki/Place\\_of\\_articulation](http://en.wikipedia.org/wiki/Place_of_articulation)