A Research Study of the Development of a Computer Aided/Assisted Instruction (CAI) System for the Japanese Language Education Based on a Spoken Dialogue Model

Masatake DANTSUJI
Faculty of Letters, Kansai University
3-3-35 Yamate-cho, Suita-shi, Osaka 564, JAPAN
e-mail: dantsuji@mm01.mm.kansai-u.ac.jp
Taizo UMEZAKI
Faculty of Engineering, Chubu University
1200 Matsumoto-cho, Kasugai-shi, Aichi 487, Japan
e-mail: ume@solan.chubu.ac.jp

The aim of this study is to construct and develop a computer aided/assisted instruction (CAI) system for the elementary course of the Japanese language education. This system is developed for the purpose of helping to reduce and to save labor of language teachers. This system is designed to make it possible to carry out repeated practices of corrected pronunciation of target words and repeated exercises of self-teachings of the elementary course of Japanese by students based on a spoken dialogue model. We carried forward our study by uniting a knowledge of acoustic phonetics and that of speech processing and speech recognition technology.

It has been pointed out that the field of the Japanese language education has not utilized the advanced technology and knowledge of speech processing adequately so far. On the other hand, engineers of the information science have seldom had opportunity to take part in the practical language education. Therefore, it is one of the purposes of this study that the researchers of both these fields combined their efforts to make progress in developing the computer aided/assisted instruction system for the language education.

In order to develop such a kind of computer aided/assisted instruction system, we have carried out fundamental research studies as follows: 1)Construction of purpose oriented tasks based on a spoken dialogue model. 2)Collecting, editing and analyzing of the speech sounds based on our spoken dialogue model. 3)Revaluation of basic and fundamental vocabularies and lexical items which this system should contain in a built-in spoken dialogue database. 4)Construction of a phonological lexicon/dictionary for the elementary course of the Japanese language education. 5)Construction of a man-machine interface which controls input devices such as speech sounds, mouse and keyboard.

One of the research items of this study is to construct purpose-oriented tasks of spoken dialogues. This study intends to realize spoken dialogues between men and machines. As one of the pilot studies, we simulate the situation of the daily life of foreign students in and around universities. It is necessary for the fundamental research of this kind of study to collect and analyze spoken dialogue arising in real situations of the daily life in and around universities. We have recorded and analyzed the spoken dialogue collected at the university libraries. We have also collected and edited materials of the spoken dialogue with respect to graduate theses of the university, too. From the results of these experiments, it was shown that a small number of vocabulary was able to cover whole the dialogue efficiently in case that we set up purpose-oriented tasks.

We have examined articulatory properties of the pronunciation of the Japanese language by applying the procedures of the experimental phonetics. The results were processed into computer by making use of both of the developed software for this system and existing software. We have extracted acoustic parameters such as formant structures from the inputted speech data of the language learners. We make it possible for learners to understand their faults of pronunciation audio visually by means of computer graphics such as vowel charts and the outputted model voice. On the computer monitor screen speech waves as well as spectrum envelope, fundamental frequency contours and so on are displayed to help teachers to evaluate the pronunciation of their students.

In research study of the construction of a phonological lexicon, we have proposed a lexicon which includes multiple layers of description such as Japanese syllabary layer, phonemic layer, allophonic or phonetic layer and the layer of distinctive features. In relation to the construction of a phonological lexicon, we have pointed out a defect of the current system of the computer coding of the international phonetic alphabet (IPA). Instead of the current coding system, we have proposed a new computer coding system of the international phonetic alphabet based on phonetic category units such as phonation (voiced vs. voiceless), places of articulation (bilabial, labiodental, dental, alveolar, postalveolar, retroflex, palatal, velar, uvular, pharyngeal and glottal), and manners of articulation (plosive, nasal, trill, tap or flap, fricative, lateral fricative, approximant, lateral approximant, click, implosive and ejective).

Examples are as follows: Voiceless sounds=1nn, Voiced sounds=2nn; Bilabials=n1n, Labiodentals=n2n, Dentals=n3n, Alveolars=n4n, Postalveolars=n5n, Retroflexes=n6n, Palatals=n7n, Velars=n8n, Uvulars=n9n, Pharyngeals=nAn, Glottals=nBn; Plosives=nn1, Nasals=nn2, Trills=nn3, Taps or Flaps=nn4, Frica- tives=nn5, Lateral fricatives=nn6, Approximants=nn7, Lateral approximants=nn8, Clicks=nn9, (Voiced) Implosives=nnA, Ejectives=nnB. Voiceless Bilabial Plosive [p]=111, Voiced Bilabial Plosive [b]=211, Voiceless Alveolar Plosive [t]=141, Voiced Alveolar Plosive [d]=241, Voiceless Velar Plosive [k]=181, Voiced Velar Plosive [g]=281, (Voiced) Bilabial Nasal [m]=212, (Voiced) Alveolar Nasal [n]=242, (Voiced) Alveolar Trill [r]=243, Voiceless Alveolar Fricative [s]=145, Voiced Alveolar Fricative [z]=245, Voiceless Glottal Fricative [h] =1B5.

In relation to the construction of a man-machine interface, we have developed a voice analytic display system by making use of graphical user interface (GUI) on X Window System that is a standard window system with respect to current work stations. This voice analytic display system is expected to show effective results at a pronunciation training process of this computer aided/assisted instruction system for the elementary course of the Japanese language education.

We have designed and constructed classrooms called "LL (language laboratory) which are equipped with personal computers for speech processing. These classes have 46 booths and additional 2 booths for spare. Each booth has a cassette deck and a headphone set with a microphone. Two booths share one built-in monitor with each other. There is a 100-inch screen in the front of the classroom. It is possible to display not only same pictures but also different pictures between booth monitors and the front screen.

Our language education which uses these rooms is oriented to utilize spoken dialogues and purpose oriented tasks such as "At the Customs", "Checkin at a hotel", "Ordering at a restaurant" and so on. Every student is assigned some role and he/she has to play his/her part of role during the lesson. Model voices of native speakers are given at first and every student has to repeat target sentences following to the model voices. After achieving this practice, he/she begins to carry out in practice what he/she has learned in spoken dialogue lessons. Each spoken dialogue of each pair can be inputted and stored into the personal computers at a master console in the front of the classroom. Each spoken dialogue can be analyzed acoustically and displayed at both booth monitors and the front screen. Therefore, every student can understand the formant structures, fundamental frequencies, durational patterns and so on of a given spoken dialogue.

Keywords: CAI (Computer aided/assisted instruction), dialogue modeling, IPA (the International Phonetic Alphabet), GUI (graphical user interface), LL (language laboratory)