DEVELOPMENT OF ARTICULATION ASSESSMENT AND TRAINING SYSTEM WITH SPEECH RECOGNITION AND ARTICULATION TRAINING STRATEGIES SELECTION

Yeou-Jiunn Chen[1], Jing-Wei Huang[1], Hui-Mei Yang[2], Yi-Hui Lin[2], and Jiunn-Liang Wu[2]

[1]Department of Electrical Engineering,
Southern Taiwan University of Technology, Tainan County, Taiwan, R.O.C.
[2]Department of Otolaryngology, National Cheng Kung University Hospital, Tainan, Taiwan, R.O.C.

ABSTRACT

Articulation problems seriously reduce speech intelligent and speech communication and affect person's interpersonal communication, personality, social adaptive capacity, and learning ability. In the clinical protocol, language therapist utilizes clinical experience to individualized assessment, treatment, and training. However, the manpower of language therapists and the assistant instruments are insufficient. In this paper, an articulation assessment and training system is proposed to assist language therapists and articulation disorders. The articulation errors in phonetic are analyzed and modeled by clinical linguist. Using clinical experience of language therapists, articulation training strategies for each type of articulation errors are designed. The articulation characteristics of user can be effectively detected. Speechreading information is responded to improve the performance of training program. The articulation training strategy is automatically selected to suggest articulation disorder in language training activities. Experimental results reveal the practicability of proposed method and system.

Index Terms **—** Articulation disorder, articulation assessment, articulation training, training strategies selection, speech recognition

1. INTRODUCTION

Articulation problems, which generate different degrees of abnormality in articulation, seriously reduce speech intelligent and speech communication. It also affects person's interpersonal communication, personality, social adaptive capacity, and learning ability. In Taiwan, the rate of language disorder is about 10% [1]. The rate of language disorder, which is aged from 4 years old to 15 years old, is 2.64%; amount them the rate of articulation disorder is 43.36% [2]. However, the population of language disorder is about 2.5 million and the linguistic therapist worked in hospital is about three hundred people [3]. It is clear that the manpower of linguistic therapist is insufficient.

Recently, clinical researches had been focused on the design of articulation test and materials [4-7]. In the sequence of therapy, language therapist subjectively utilizes clinical experience to individualized assessment, treatment, and training [7-9]. In Taiwan, Augmentative and Alternative Communication had been supported by National Science Council and Ministry of the Interior and used to help language disorder [10]. However, most of assistive devices are used for alternative communication not for articulation assessment and training. The assistant instruments used by language therapist can support acoustic information such as spectrum information, energy, pitch, etc. Hence, articulation assessment and training system is the imperious demands for language therapist and articulation disorder.

For the articulation assessment, Glassman applied spectral parameters such as filter bank information to separate fourteen pair of consonants [11]. Neural network had been used to detect the articulation error types [12]. However, the articulation errors in phonetics are laminated in several articulation errors and cannot support complete articulation information. For the articulation training, language therapists have to individually assess and train articulation ability. The training program is processed by articulation disorder at home. Therefore, assistant instructions for articulation training are very important and insufficient.

In this paper, an articulation assessment and training system is proposed to help language therapists and articulation disorders. Speech recognition techniques are adopted to automatically assess the articulation errors in phonetics. Speechreading information and articulation training strategy can be replied to user. Using language training activities, customary articulation ability can be treated in games.

2. MATERIALS AND METHODS

Fig. 1 shows the proposed framework of articulation assessment and training system. First, the testing vocabularies are designed by language therapist and used to assess articulation errors of articulation disorder. In order to improve the performance of articulation training program, speechreading responses of testing vocabularies and language training activities of articulation errors are designed and recorded. In training program, a prompt is shown on the screen and the user is asked to pronounce it. Speech recognition is applied to assess the articulation

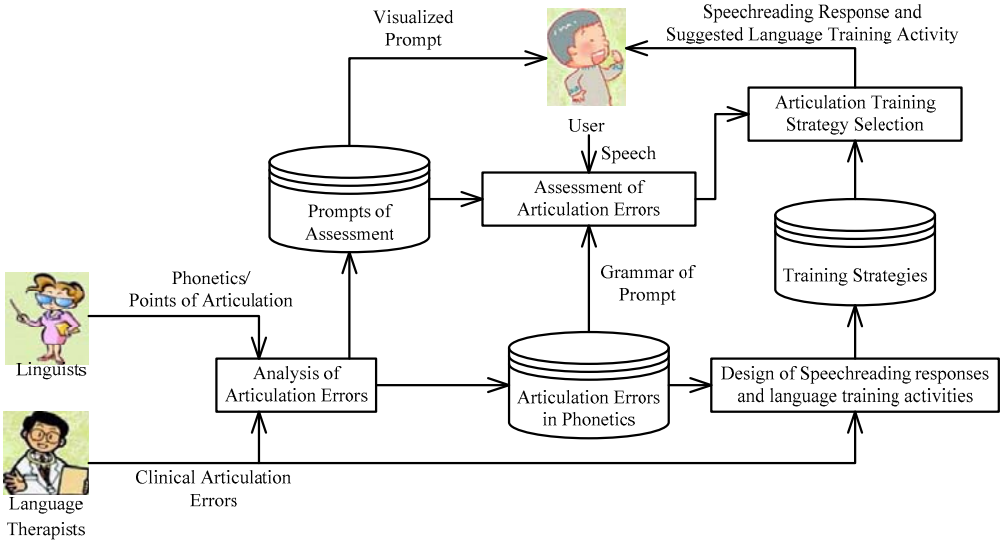

Fig. 1 The framework of articulation assessment and training system

characteristic. Speechreading responses and language training activities in games are automatically selected to assist articulation disorders in practicing articulation.

2.1. Analysis of articulation errors

In Mandarin speech, all syllables are monosyllabic and each syllable can be phonetically decomposed into two parts: an INITIAL and a FINAL. The INITIAL of a syllable is optional and comprises a single consonant if it exists. The FINAL comprises a vowel or diphthong nucleus preceded by an optional medial and followed by an optional nasal. The relationship between points of articulation and INITIALs/FINALs is analyzed by linguists and shown in Table 1 [13]. In clinical protocol, articulation disorder can be categorized as (1) substitution errors: fronting, backing, unaspirated, stopping, and palatalization; (2) deletion errors: consonant deletion, optional medial deletion, and nasal deletion; (3) others: retroflexion deletion et al. [2, 14]. Using clinical experience, the articulation errors can be detected from Table 1 and modeled as replication rules shown as Table 2. Moreover, the prompts used to assess the articulation ability are also designed by linguists and language therapist with two selection criterions: high familiarity and phonetic balance. A finite state grammar for a prompt is also generated by the replication rules. An example of prompt "qi shui" is shown in Fig. 2.

2.2. Design of training strategies

In the clinical protocol, the sequence of treatments for articulation disorder is: (1) establishment of self articulation problem consciousness; (2) discriminative of correct or incorrect articulation; (3) establishment of correct articulation; (4) language therapy/training in games [15, 16]. Therefore, five training strategies named articulation error awareness, speech discrimination training, provoking articulation practice, articulation habit treatment, and sentence based articulation training are designed for each type of articulation errors. The results of language training activities are shown in Table 3. Moreover, a speechreading based feedback is responded immediately to user to excite user's articulation.

2.3. Assessment of articulation errors

When a visualized prompt is shown on the screen, the speech signal pronounced by user can be captured and parameterized as observation sequence, $O=o_1o_2...o_T$. T is the total number frames of input speech. Furthermore, the corresponding grammar of prompt W_p is also adopted to reduce search space of Viterbi algorithm and find the most likely prompt W_l, where

$$W_l = \arg\max_i L\left(O|W_i\right) \tag{1}$$

and $L(O|W_i)$ is the likelihood of the observation sequence O given prompt W_i [17]. The prompt can be treated as a sequence of subsyllable units in Mandarin and shown as

$$W_i = s_1^i s_2^i \cdots s_N^i \tag{2}$$

where N is the total number of subsyllables. For each subsyllable s, a Hidden Markov Model (HMM), $\lambda_s=(A_s, B_s, \pi_s)$, is established and estimated to optimize the likelihood of the training set of subsyllable s. A_s, B_s, and π_s are the state transition probability distribution, observation symbol probability distribution, and initial state distribution, respectively. Comparing W_p and W_l, the articulation errors in phonetics can be detected.

Table 1. The relationship between points of articulation and INITIALs/FINALs

(a) INITIALs

	Stops		Affricate		Fricative	Nasal	Lateral	Trill
	Unaspirated	Aspirated	Unaspirated	Aspirated				
Bilabial	b	p				m		
Labiodental					f			
Dentoalveolar			zi	ci	si			
Alveolar	d	t				n	l	
Retroflex			zhi	chi	shi			ri
Alveopalatal			j	q	x			
Velar	g	k			h			

(b) FINALs

	kai kou hu	qi chi hu	he kou hu	cuo kou hu
Mono-Vowel		yi	wu	yu
	a	ya	wa	
	o	yo	wo	
	e			
	e	ye		yue
	er			
Composed-Vowel	ai	yai	wai	
	ei		wei	
	ao	yao		
	ou	you		
Nasal-Vowel	an	yan	wan	yuan
	en	yin	wen	yun
	ang	yang	wang	
	eng	ying	weng	yung

2.4. Articulation training strategy selection

The relationship of articulation error types between each phoneme is analyzed by language therapists and shown in Table 3. The number in Table 3 is the no of articulation errors. Integrating Table 3 and detected articulation errors, the probability for each type of articulation errors can be estimated. Using statistical process, a threshold can be objectively selected by language therapist and used to eliminate some incorrect decisions generated by speech recognition errors. According to articulation characteristics, the language training activities in games are designed by language therapists and used help articulation disorders to correct articulation habit. Some language training activities in games are shown in Table 4.

4. EXPERIMENTAL RESULTS

For the parameters of speech signal, a 26-dimension feature vector, including 12 MFCCs, 12 delta MFCCs, delta log energy, and delta delta log energy was extracted. 94 right context-dependent INITIAL and 38 context-independent FINAL HMMs were constructed in assessment of

Table 2. Replication rules

INITIALs		FINALs	
Target Phonemes	Replication Phonemes	Target Phonemes	Replication Phonemes
p	b	ying, yin	yi
chi, shi	zhi	yuan	yua
q, x	shi	en, eng	e
g, k, zi, ci, zhi, q	d, t	wa	a
zi, ci, zhi, chi, si, shi	g, k	wei	ei

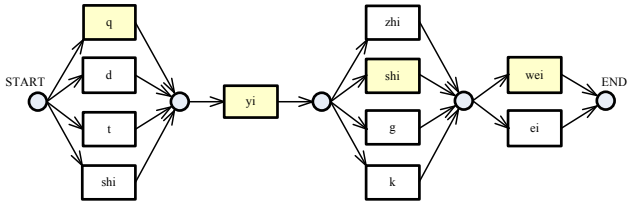

Fig. 2. Example of grammar for prompt “qi shui”

Table 3. The relationship of articulation error types between each phoneme

Target \ Pronounced	b	p	m	f	d	t
b						
p	3					
m	4	4			4	4
f	4	4			4	4
d						

Table 4. Part of language training activities in games

No	Language training activities
1	Utilize action of the cough, child is taught what the throat sound is.
2	Utilize the action of blow, child is taught what aspiration is.
3	Utilize the enunciation simulation, child is taught what alveolar is.
4	Pronounce the long sound x~~yu ,It is fricative.

articulation errors. The HMM of INITIALs and FINALs consists 3 states and 5 states, each with 10 Gaussian mixture densities.

2676 sentences (29 males, 25 females) of TCC300 were used as the training corpus and used to estimate the parameters of HMMs. TCC300 was recorded by microphone in Taiwan and the recording texts were selected from "Academia Sinica Balanced Corpus" [18]. In addition, 2412 and 1482 sentences were also recorded by 39 children, whose average age is 5, for correct or incorrect articulation, respectively. A questionnaire was designed to investigate user's satisfaction at system interface. The users give mean opinion scores (MOS) on a scale of 1 to 5, i.e., 5 for excellent level, 4 for good level, 3 for fair level, 2 for poor level, and 1 for unsatisfactory level. The average MOS was 4.13. Articulation detection rate in sentences could achieve 82.7%. Although the detection performance is not perfect, articulation training strategy selection tends to find consistent in articulation errors and the effects of detection error can be tolerant.

5. CONCLUSIONS

An articulation assessment and training system was proposed to assist language therapists and articulation disorders in this paper. Utilize the analysis of articulation errors in phonetics; the search space of speech recognition can be effectively reduced. Thus, speech recognition can be successfully applied to detect articulation errors in phonetics. With the articulation detection mechanism, articulation disorder can assess articulation characteristic by him at home. The feedback responses language training activities can effectively improve the efficiency of articulation training. Experimental results show the proposed system is very promising in articulation training.

6. ACKNOWLEDGEMENT

The authors would like to thank the National Science Council, R.O.C., for its financial support of this work, under Contract No. NSC95-2221-E-218-002-MY2.

7. REFERENCES

[1] Available: http://www.moi.gov.tw/stat/english/index.asp Department of Statistic, Ministry of the Interior.

[2] B.G. Lin, *Language disorder and treatment*, Wunan Book Co., Ltd.1994.

[3] H. Sheng et al, *Report of the police recommendation for manpower of linguistic therapist*, The Magazine of Hearing and Language, The Speech-Language-Hearing Association of the Republic of China, vol. 16, pp 76-91, 2001.

[4] M.N. Hegde, *Introduction to communication disorders*, Austin, Texas: PRO-ED, 1991.

[5] B.G. Lin and X.M. Lin, *Appraisal of language disorder for pre-school children*, 1993.

[6] H.M. Yang, J.L. Wu, S.L. Lin, Y.H. Lin, "*Mandarin monosyllabic lexical neighborhood nest: Interlist equivalancy and lexical effect*," Cochlear Implant Int, pp. 201-203, 2004.

[7] G.J. My, "*The strategy of language evaluation and treatment for children with language and articulation disorder*," The Speech-Language-Hearing Association of the Republic of China, pp 29-35, 1998.

[8] X.J. Lai, "*Articulation disorder*," The Magazine of Hearing and Language, The Speech-Language-Hearing Association of the Republic of China, vol. 4, pp 70-73, 1987.

[9] C. Van Riper and L. Emerick, *Speech correction: An introduction to special pathology and audiology*, N.J.: Prentice-Hall, 1984.

[10] Available: http://ciat.moi.gov.tw Center for Information Assistive Technology, Ministry of the Interior.

[11] M.S. Glassman and M.B. Starkey, "*Speech therapy using computer based minimal consonant pair discrimination*," IEEE Proc, Engineering in Medicine and Biology Society, pp 1421-1422, 1988.

[12] Y.Q. Chen, *Apply neural network to detect articulation error type of children, Master Thesis*, Department of Engineering Science, National Cheng Kung University, 1995.

[13] G.P. Sie, *Introduction to linguistics*, SanMin Publishing, 2006.

[14] S.J. Lai, "*Diagnostic methods of speech and management*," The Magazine of Language Therapy, Taipei city Government Educational site, pp.123-133, 1990.

[15] B.L. Yu, "*Appraisal and therapy of language disorder for children*," The Speech-Language-Hearing Association of The Republic of China, pp.29-35, 1992.

[16] S.L. Wu, *Teaching activities of language disorder,* National Kaohsiung Normal University Special Education Center, 2000.

[17] L. Rabiner, *Fundamentals of speech recognition*, Prentice Hall PTR, 1993.

[18] Available: http://www.aclclp.org.tw The Association for Computational Linguistics and Chinese Language Processing