Malay Word Pronunciation Test Application for Pre-School Children

M.Y. Shahrul Azmi,

Abstract—In Malaysia, many researchers focus on developing independent speaker speech recognition systems that uses Malay Language. Accuracy, noise robustness and processing time are concerns when developing speech therapy systems especially for children. In this study, a Malay word pronunciation test application is developed using Spectrum Delta features and Logistic Regression classification model in an effort to improve Malay word pronunciation for pre-school children aged between 3-6 years old. Results showed that the pronunciation application can assist children to test and improve their Malay word pronunciation.

Index Terms—Malay Word, Pronunciation Test, Pre-School Children

1 Introduction

Computer based speech therapy and assessment is still new in Malaysia. In Bahasa Malaysia language, children are normally taught to spell Malay words using a combination of consonant and vowel sounds such as "BACA" represented by syllable "BA" and "CA". Speech therapy that uses vowel phonemes can be used to improve Malay word pronunciation for children. A hearing impaired person can be trained to speak properly with a good degree of intelligibility in pronouncing words. A high degree of standard Malay vowel recognition capability is needed in all of these systems.

Although there are many studies on Malay phoneme recognition, more work still needs to be done. Most of these studies use multiple frame analysis. Accuracy, noise robustness and processing time are still concerns when developing speech therapy systems, especially for children using Malay Language. The accuracy aspect involves factors such as age and gender. The size of the vocal tract of children of different gender and age varies which causes their voice to have different fundamental frequencies. Any application that uses vowel phonemes requires a high degree of Standard Malay (SM) vowel recognition capability. This motivates this study to have an objective of developing a Malay word pronunciation test application in an effort to improve Malay word pronunciation for pre-school children aged between 3-6 years old.

2 Malay Speech Therapy Systems

A Computer-based Malay Language Articulation Diagnostic System was developed using Hidden Markov Model (HMM) and Mel-Frequency Cepstral Coefficients (MFCCs) [1]. It was developed using a database of Malay words. In 2012, Tan et.al developed a Malay dialect translation and synthesis system, but still at a preliminary stage [2]. The speech synthesis system used here is an HMM speech synthesis system (HTS Speech Synthesis System) at a sampling rate of 22 kHz. The results were promising, but the system does not test on pronunciation. A research was done in 2014 with the objective of developing an ASR system for Malay-speaking children [3]. The speech corpus comprises of six children uttering a total of 390 sentences. The parameter training is performed using the HTK toolkit by utilizing an HMM speech acoustic model of Malay-speaking children. The system can accurately recognize of up to 76% of test words. Yusof et.al did a study about speech intelligibility of deaf children in Malaysia using a Malay Speech Intelligibility Test (MSIT) system [4]. Researchers from Universiti Malaysia Sarawak did a study on syllabification algorithm based on Malay syllable structure [5]. It was used to build the Iban and Bidayuh syllable list and speech corpus. The accuracy, using Categorical Estimation (CE), gave a mean score of 3.07 out of 5.

3 Malay Word Pronunciation Application

2.1 Speech Recognition Engine

The Malay Word Pronunciation Application engine is based on vowel recognition process. It started with the data acquisition, followed by filtering and pre-processing, frame selection, speech signal modelling, feature extraction and vowel recognition processes as shown in Fig.1.

• *M.Y. Shahrul Azmi is with the UUM Technopreneur Incubation Center Universiti Utara Malaysia, Malaysia, Kedah 06010. E-mail: shahrulazmi@uum.edu.my*

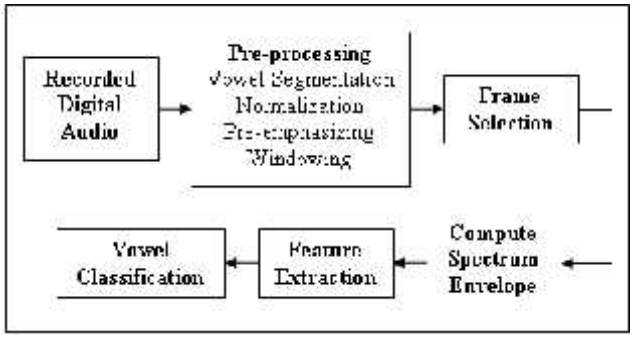

Fig 1. Vowel recognition process

2.2 Data Acquisition

The data collection process was done and taken from 20 Malay pre-school children from Changlun and Jitra area in Kedah. The words "ka", "ke", "ki", "ko", "ku" and "kə" were recorded from speakers representing six vowels of /a/, /e/, /i/, /o/, /u/ and /ə/. In this study, 8000 Hz sampling frequency was used to sample the vowels and up to 10 recordings were taken per speaker depending on situation convenience.

2.3 Spectrum Delta Feature Extraction Method

The speech feature used in this study is Spectrum Delta developed by Shahrul Azmi [6]. In this Spectral Delta approach, the differences in amplitude between each frequency band were used as features. First, the band where most of the vowel energy is situated is divided into three frequency band regions. The frequency of interest is between 1 to 2350Hz and divided into three equal band regions of 780Hz.

2.4 Logistic Regression

Many researchers have used logistic regression technique in speech recognition to classify words [7], [8]. Logistic regression is a type of category of statistical models and is commonly called generalized linear models. It can be used to predict the absence or presence of an outcome based on a set of predictor variable values. Binomial logistic regression is used when the dependent variable is a dichotomy or having 2 possible values and the independent variable can be continuous variables or categorical variables, or even both. Multinomial logistic regression (MLR) handles the case of dependent variables with more than 3 levels. The details of this technique can be referred from [9].

2.5 Pronunciation Interface

The interface was developed using Matlab. The words "katil", "roti", "lara", "buncis", "potong" and "betik" were selected representing the six vowels of /a/, /e/, /i/, /o/, /u/ and /ə/. The interface is as shown in Fig. 2. The lowest accuracy was recorded based on the first utterance of word. Then the children are trained on how to pronounce the words properly. Next the children will try uttering the words again 4 times and the best accuracy result is then taken.

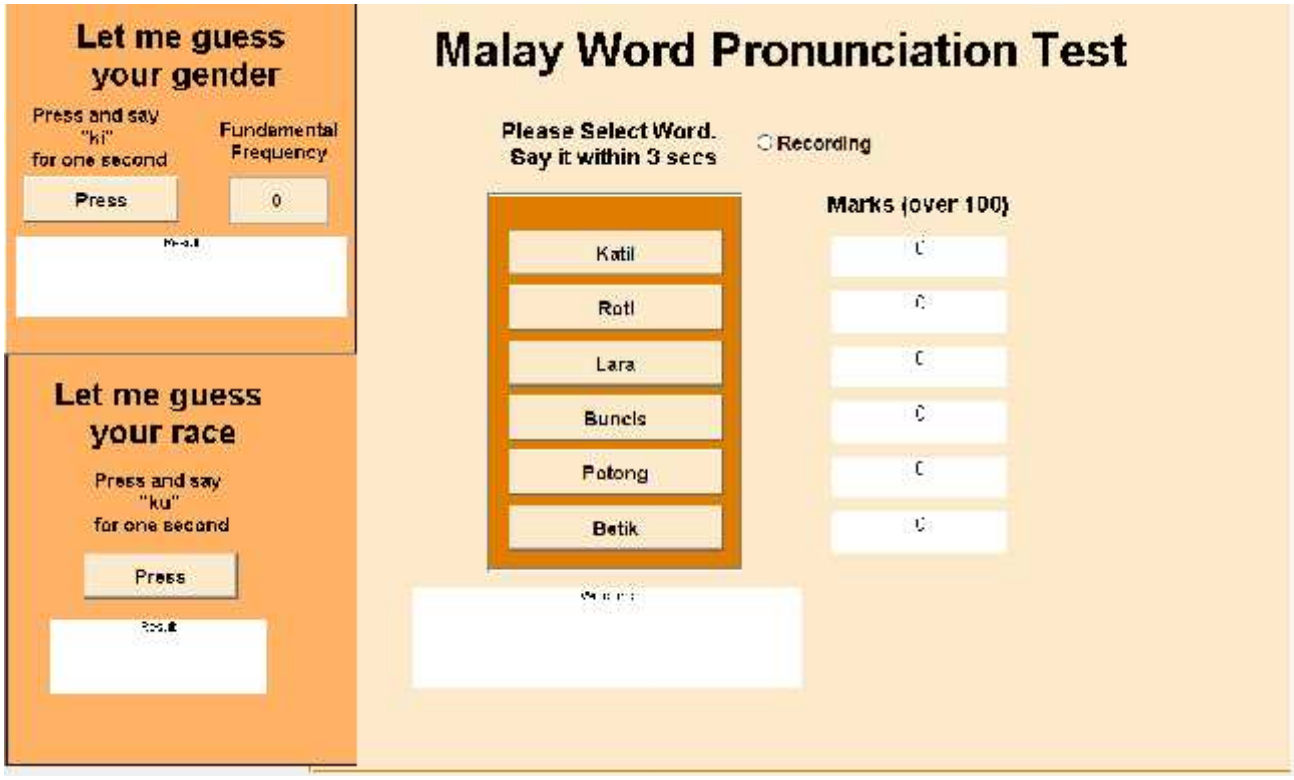

Fig 2. Matlab Screenshot of Testing Interface

2.6 Results and Discussion

The application is tested on 10 Malay pre-school children aged between 3 to 6 years from several kindergartens in Changlun and Sintok as shown in Fig. 2. The results of the pronunciation test are shown in Table 1. The average lowest accuracy is the average accuracy of the first trial of each speaker pronunciation the given words.

Fig 2. Data Collection Process

TABLE 1. PRONUNCIATION TEST RESULTS

	Accuracy		
Words	Average Lowest	Average Highest	Average Tries to reach average highest
Katil	56%	86%	4
Roti	65%	92%	4
Lara	71%	92%	2
Buncis	52%	84%	4
Potong	70%	95%	2
Betik	58%	88%	3

The first accuracy results were taken based on the first attempt. Then the speaker will be trained on how to pronounce properly and tested again in the next 5 more attempts. Based on the results obtained, for the word "katil" or "Ka" and "til" meaning bed, the average lowest accuracy were 56% accuracy for the first trial and 86% for after improvements were done in subsequent pronunciations. On the average, 4 times are needed to obtain the highest average accuracy. For this word, the syllable "til" was supposed to be pronounced like the word "till" in English but often pronounced inaccurately as "tail" in English. After the correction in pronunciation, the children improve their pronunciation and obtained an improved score of 86%. For the word "Roti", the children often mispronounced "Ro" as "Rue" instead of "Roo". For the word Lara, not much problem in pronouncing it correctly, but the lower initial accuracy was due to the speaker spoke the word softly and sounded less confident. For the word "Buncis", the syllable "cis" was often mispronounced as "cess" instead of "cheese" in English. Not much problem seen when pronouncing the word "potong". For the word "Betik", the syllable "tik" was often mispronounced as "take" instead of "tick" in English. It would seem that the tested children often mispronounced syllables having vowels /i/, /e/, /o/ and /u/. This is due to the common pronunciation they normally use during normal conversation. Overall, this application is able to differentiate the differences in terms of syllable pronunciations and able to give scores corresponding to the words being uttered.

It was also observed that the children who have louder voices tend to speak clearly with a higher degree of speech intelligibility than the soft-spoken children. This was evident in the higher scores obtained by them and they tend to reach maximum accuracy scores faster than soft spoken children. This was mostly due to shyness, nervousness and a bit lack of confidence.

3 CONCLUSION

This paper presents a new Malay Word Pronunciation Application for pre-school children. This application was developed using Matlab and uses a speech recognition algorithm based on Spectrum Delta features and logistic regression classification technique. The application was developed and tested on pre-school children aged between 3 to 6 years old. Based on this study, the vowels /i/, /e/, /o/ and /u/ are often mispronounced. This is because of the lack of emphasizing on proper pronunciation of the given words due to daily mispronunciation which is often happening around the children. The clarity of the pronounced words may lower the accuracy measured by the application which is mostly due to nervousness and lack of confidence. Overall, this application may able to help pre-school children learn to pronounce Malay words properly and clearly.

Acknowledgments

The author would like to thank Universiti Utara Malaysia for providing the grant to do this study. The author would

also like to appreciate all the colleagues and students who have supported in giving good insights and suggestions.

REFERENCES

[1] Mazenan, M. N., & Tan, T. S. (2013). "Malay Alveolar Vocabulary Design for Malay Speech Therapy System". *Recent Advances in Electrical Engineering Series*, (11).

[2] Tan, T. P., Goh, S. S., & Khaw, Y. M. (2012, November). "A Malay Dialect Translation and Synthesis System: Proposal and Preliminary System". Paper presented at the *IEEE Asian Language Processing (IALP), 2012 International Conference on* (pp. 109-112).

[3] Rahman, F. D., Mohamed, N., Mustafa, M. B., & Salim, S. S. (2014, March). "Automatic speech recognition system for Malay speaking children". Paper presented at the IEEE *Student Project Conference (ICT-ISPC), 2014 Third ICT International* (pp. 79-82).

[4] Yusof, Z. M., Hussain, R., & Ahmed, M. (2014). "Malay Speech Intelligibility Test (MSIT) for Deaf Malaysian Children". *International Journal of Integrated Engineering*, *5*(3).

[5] Juan, S. F., Edwin, V., Cheong, C. Y., Choi, L. J., & Yeo, A. W. (2011, November). "Adopting Malay Syllable Structure for Syllable Based Speech Synthesizer for Iban and Bidayuh Languages". In *Asian Language Processing (IALP), 2011 International Conference on* (pp. 216-219). IEEE.

[6] Shahrul Azmi, M., Nor Idayu, M., Roshidi, D., Yaakob, A. R., & Yaacob, S. "Noise Robustness of Spectrum Delta (SpD) Features in Malay Vowel Recognition". In *Computer Applications for Communication, Networking, and Digital Contents* (Vol. 350, pp. 270-277): Springer Berlin Heidelberg.

[7] Birkenes, Ø. (2007). "A Framework for Speech Recognition using Logistic Regression". Unpublished PhD thesis, Norwegian University of Science and Technology.

[8] Hillenbrand, J., & Nearey, T. (1999). "Identification of resynthesized/hVd/utterances: Effects of formant contour". *The Journal of the Acoustical Society of America, 105*(6), 3509-3523.

[9] So, Y., & Cary, N. (1995). "A tutorial on logistic regression". *SAS White Papers*.

M.Y. Shahrul Azmi obtained his bachelor degree in Electrical Engineering from Missouri University of Science and Technology, USA (previously University of Missouri-Rolla) in 1994. He later pursue his MSc. In Information Technology from Universiti Sains Malaysia and obtained his PhD in Mechatronics Engineering from Universiti Malaysia Perlis in 2010. He served as an engineer at Robert Bosch and Seagate Company from 1995 until 1999. Currently, he is a faculty member at School of Computing in Universiti Utara Malaysia. His research interests include Speech Recognition, IT Entrepreneurship, Artificial Intelligence and Manufacturing.