

**THE PRONUNCIATION ACCURACY OF INTERACTIVE
DIALOG SYSTEM FOR MALAYSIAN PRIMARY SCHOOL
STUDENTS**

Hasliemelia Binti Abu Hassan @ Aziz

University Utara Malaysia

2012

**THE PRONUNCIATION ACCURACY OF INTERACTIVE
DIALOG SYSTEM FOR MALAYSIAN PRIMARY SCHOOL
STUDENTS**

A project submitted to the Graduate School in partial
Fulfillment of the requirements for the degree of
Master of Science (Information Technology)
University Utara Malaysia

By

Hasliemelia Binti Abu Hassan @ Aziz (805737)

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from University Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence by the Dean of the Postgraduate and Research.

It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to University Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part should be addressed to:

Dean of Research and Postgraduate Studies

Collage of Arts and Sciences

University Utara Malaysia

06010 UUM Sintok

Kedah Darul Aman

Malaysia

ABSTRAK

Projek ini adalah untuk mengkaji ketepatan pengecaman suara dalam sebutan Bahasa Inggeris menggunakan sistem yang dibina. Penggunaan enjin pengecaman suara yang sedia ada dalam sistem dialog interaktif, untuk digunakan oleh pelajar sekolah rendah sebagai bahasa kedua di Malaysia dalam literasi pendidikan. Ia dapat memupuk minat pelajar menggunakan ICT dalam pembelajaran serta mendorong pelajar untuk menjadi lebih yakin dalam sebutan dan bacaan tanpa bantuan dari guru semata-mata. Pengajaran dan pembelajaran menggunakan komputer ini juga dapat meningkatkan tahap keupayaan pelajar belajar membaca dengan sebutan yang betul dan tepat kerana sistem ini dapat mengenalpasti sebutan yang tepat atau salah. Pembelajaran menggunakan komputer ini juga akan meningkatkan keupayaan pelajar membaca lisan dengan menggunakan sistem ini di dalam komputer. Kajian ini dijalankan di Sekolah Kebangsaan Sungai Berembang yang melibatkan 16 orang pelajar perempuan dan 18 orang pelajar lelaki tahun dua yang berusia lapan tahun. Sudah tentunya pelajar-pelajar ini mempunyai pelbagai pelat sebutan bacaan, kebolehan dan pengalaman dalam bahasa Inggeris kerana bahasa ibunda mereka adalah bahasa Melayu. Objektif utama kajian ini ialah untuk mengenalpasti ketepatan menggunakan sistem pengecaman suara yang sedia ada bagi pelajar yang menggunakan Bahasa Inggeris sebagai bahasa kedua dalam literasi pendidikan di Malaysia. Objektif khusus kajian ini adalah untuk mengenalpasti keperluan penggunaan sistem pengecaman suara dan menilai ucapan dialog menggunakan sistem tersebut berasaskan kejituan bacaan. Teknologi pengecaman suara ini bertujuan untuk membantu guru dalam menjalankan pengajaran dan pembelajaran di sekolah agar dapat meningkatkan keupayaan dalam kesedaran fonemik kanak-kanak, pembangunan perbendaharaan kata, pemahaman perkataan, dan pembacaan yang lancar. Kaedah yang digunakan terbahagi kepada lima peringkat seperti berikut iaitu membina Kerangka Membangunkan Senibina Sistem, Analisa dan Rekabentuk Sistem, Membina Sistem (prototaip) dan Pemerhatian semasa menguji sistem. Hasil daripada kajian dan pelaksanaan terhadap pelajar didapati 85% IDS ini berjaya membantu dengan berkesan penguasaan sebutan Bahasa Inggeris selepas menggunakan sistem ini.

ABSTRACT

This project is to examine the accuracy of using existing speech recognition engine in interactive dialog system for English as second language (ESL) Malaysian primary school student in literacy education. Students are interested to learn literacy using computer that encompasses spoken dialog as it motivates students to be more confidence in reading and pronunciation without depending solely on teachers. This computer assisted learning will improve student's oral reading ability by using the speech recognition in IDS. By using the system students are able to learn, to read and pronounce a word correctly independently without seeking help from teachers. This study is conducted at Sungai Berembang Primary School involving all 16 female and 18 male standard 2 students aged 8 years old. These students possess various reading pronunciation, abilities, and experience in English language with Malay language as their first language. The main objective of this study is to examine the accuracy of using an existing speech recognition engine for ESL Malaysian students in literacy education. The specific objectives of this study are to identify requirement and evaluate speech recognition based dialog system for reading accuracy. This kind of speech recognition technology is aiming to provide teacher-similar tutoring ability in children's phonemic awareness, vocabulary building, word comprehension, and fluent reading. This method has five stages. This method enables to construct a framework. Develop system architecture then analyze and design the system. It also builds the prototype for the system upon the system implementation which will be used in this study is the System Development Research Method. Lastly its observe, test the system and the results of the study and implementation of IDS students found 85% of this has helped the English language after using this system.

DEDICATION

I humbly thank **Allah** Almighty, the Merciful and the Beneficent, who gave me health, thoughts and co-operative people to enable me achieve this goal.

I wish to dedicate this work to **Holy Prophet Muhammad** (Peace be upon him) and his companions who laid the foundations of Modern civilization and paved the way for social, moral, political, economical, cultural and physical revolution.

To my supervisors, **Madam Zahurin Bt Mat Aji @ Alon** and **Dr Husniza Bt Husni**, who always stood beside me and helped me to do my best, for the time they spent to guide me, I will always be their student throughout my life.

I would also like to say thanks to **Miss Sharmila Mat Yusof** and **Dr . Kang Eng Thye**, my project evaluators for their insightful comments. In addition, special thanks to University Utara Malaysia for giving me a chance to complete my master with great experience.

I also have a lot of thanks to my lovely husband, **Baharuddin**, my sons, **Bazli Helmy** and **Bazil Haiqal**, my daughters, **Batrisyia Ezzati**, **Balqis Hadirah**, and **Bahira Ellysa**, my mother, **Hasnah Ali** and my best friend, **Siti Fatimah** for their never ending moral supports and prayers which always acted as a catalyst in my academic life.

ACKNOWLEDGEMENT

Firstly, I thank **Allah** for helping me in my study and guiding me to continue what I have started in my educational life. I thank Allah every day for giving me the ability and motivation to continue this work.

Secondly, I would like to convey my regards to my beloved supervisors, **Madam Zahurin Binti Mat Aji @ Alon** and **Dr Husniza Binti Husni** for the provided guidance and precious information. I thank and honor their helps and support in completing my study. With their busy schedule they still spent time in guiding and sharing their.

Finally, I would like to say thankfulness word for the lecturers in the **School of Computing, University Utara Malaysia**. I would like to convey my deepest appreciation to everyone who helped me to complete my Masters Project, session 2010/2012

Thank you UUM...

Hasliemelia Binti Abu Hassan @ Aziz

2012

LIST OF TABLE

No of Table	Title	Pages
Table 1.1	Content and vocabulary in syllabus for one week	5
Table 4.1	Excerpt of a dialogue between the user (U) and the system (S).	28
Table 5.1	Shows the frequency of correct utterances recognized by the system based on each word by five students.	43
Table 5.2	Target Word	44
Table 7.1	Micro teaching	53
Table 7.2	Name for Class 2 Alpha	54

LIST OF FIGURES

No of Figure	Title	Pages
Figure 2.1	IDS are two way communication	11
Figure 2.2	The Human Dialog System	14
Figure 2.3	Interactive dialog system	15
Figure 2.4	The typical pipeline architecture of a spoken dialog system	16
Figure 3.1	System Development Research Method	19
Figure 4.1	Loading Toolkit	22
Figure 4.2	Canvas	22
Figure 4.3	RAD	23
Figure 4.4	Rename and Add Port	24
Figure 4.5	Delete Line	24
Figure 4.6	Canvas Greeting	25
Figure 4.7	Caption Greeting	25
Figure 4.8	Prompt Greeting	26
Figure 4.9	Photo picture	26
Figure 4.10	Vocabulary	29
Figure 4.11	Button Build and Run	30
Figure 4.12	RAD Flow	30
Figure 4.13	RAD prompt size_0	30
Figure 4.14	RAD for third question	31
Figure 4.15	Picture Layer of cakes	32

LIST OF FIGURES

No of Figure	Title	Pages
Figure 4.16	Window Placement for Media Object	32
Figure 4.17	Choose Face	33
Figure 4.18	Face to choose	34
Figure 4.19	Animation figure	35
Figure 4.22	Decoration of cakes	37
Figure 4.23	Picture Decoration of cakes	38
Figure 4.24	Price	39
Figure 4.25	Repair	39
Figure 4.26	Global Preferences	40
Figure 4.27	Repair default	40
Figure 4.28	Flow of all the system	41

TABLE OF CONTENTS

PERMISSION TO USE.....	I
ABSTRAK	II
ABSTRACT.....	III
DEDICATION.....	IV
ACKNOWLEDGEMENT.....	V
LIST OF TABLE	VI
LIST OF FIGURES	VII
LIST OF FIGURES	VIII
TABLE OF CONTENTS	IX
LIST OF ABBREVIATIONS	XI
CHAPTER ONE: INTRODUCTION.....	1
1.0 BACKGROUND	1
1.1 PROBLEM STATEMENT	2
1.2 PROJECT OBJECTIVE	3
1.3 SCOPE OF THE STUDY.....	3
1.4 CSLU TOOLKIT	5
1.5 SIGNIFICANCE	6
1.6 SUMMARY.....	7
CHAPTER TWO: LITERATURE REVIEW.....	8
2.0 INTRODUCTION.....	8
2.1 ASR.....	8
2.2 THE SPEECH SIGNAL AND PRODUCTION OF SPEECH.....	11
2.3 SPEECH PARAMETERS USED BY SPEECH RECOGNITION SYSTEM....	12
2.4 PREDICTING ASR PERFORMANCES USING PROSODIC CUES	13
2.5 INTERACTIVE DIALOG SYSTEM	14
2.5.1 Spoken Dialog Systems	15
2.6 READING AND CHILDREN	16
2.7 SUMMARY	17
CHAPTER THREE: METHODOLOGY	18
3.0 INTRODUCTION.....	18
3.1 CONSTRUCT FRAMEWORK	19
3.2 DEVELOP SYSTEM ARCHITECTURE	19
3.3 BUILD THE PROTOTYPE SYSTEM	20
3.3.1 RAD	20
3.3.2 RAD Development Cycle.....	20
3.4 OBSERVE AND TEST THE SYSTEM	21

3.5 SUMMARY	21
CHAPTER FOUR: IMPLEMENTATION	22
4.0 INTRODUCTION.....	22
4.1 HOW TO START	22
4.1.1 INSTRUCTIONS	23
4.2 RENAME AND ADD PORT	24
4.3 GREETING	25
4.4 PICTURE	26
4.5 VOCABULARY	29
4.6 WINDOW PLACEMENT FOR MEDIA OBJECT.....	31
4.7 CHOOSING CONVERSATIONAL AGENT.....	33
4.8 ANIMATION FIGURE	36
4.9 DECORATION OF CAKE	36
4.10 PRICE	39
4.11 REPAIR AND GLOBAL PREFERENCES.....	40
4.12 RAD EMELIA CAKE SHOP	41
CHAPTER FIVE: CONTRIBUTION AND CONCLUSION	42
5.0 INTRODUCTION.....	42
5.1 RESULT.....	42
5.2 RECOMMENDATION AND INSIGHTS	45
5.3 RESEARCH CONCLUSION.....	45
6. REFERENCES.....	48
APPENDIX.....	52
APPENDIX A	52
APPENDIX B	53
APPENDIX C	55
APPENDIX D	56
APPENDIX E	58
APPENDIX F.....	60

LIST OF ABBREVIATIONS

ASR	Automatic speech recognition
ART	Automatic Reading Tutor
ESL	English as second language
ICT	Information Communications Technology
IDS	Interactive Dialog System
TTS	Text to speech

CHAPTER ONE:INTRODUCTION

1.0 BACKGROUND

In the last few years, research in the field of interactive dialogue systems has experienced increasing growth (Pietquin & Renals, 2002). The automation of dialogue strategy design is a leading domain of investigation, and the treatment of dialogue system design using the formalism of Markov Decision Processes (MDPs) and Reinforcement Learning (RL) was proposed by Pieraccini and Levin (Choi, 2004; Levin, Pieraccini, & Eckert, 2000) .

To obtain a fully automatic procedure, the learning agent needs real interactions with a user through an automatic speech recognition (ASR) system, a large amount of corpus data or a sequence of simulated interactions with a virtual user (Dahl & Claesson, 1999). The success of a spoken dialogue system depends crucially on a carefully designed interface that can overcome the limitations of current spoken language technology (Burgt, Andernach, Kloosterman, Boston, & Nijhol, 1996; Kamm, 1995).

This system is an interactive application multimedia for ESL students. This project is about to identify requirements for automatic speech recognition and developing speech recognition based interactive dialog system for reading in English.English is one legacy of more than a century's worth of British colonial rule in Malaysia. It is the most important foreign language in Malaysia and is used extensively in practically all aspects of daily life, from conducting business transactions to labeling products to writing jingles for television advertisements. Both English and Malay language, the official language in Malaysia, play a vital role in binding together a multicultural nation made up largely of three separate and distinct races; Malay, Chinese and Indians. Even though differ in appearance and mother tongue, these groups rely on one or both languages when communicating outside their ethnic groups (in some cases even within them). Both languages help to unite people and create a unique national consciousness (Murugesan, 2003).

English language has been recognized as a global language and an international lingua franca. Despite the fact that there are more people who use English as their second or foreign language, its impact on culture and identity remains an under-researched area (Lee, Lee, & Wong, 2010). In Malaysia, English has a rather complex and ironic status. It is an “inherited”

The contents of
the thesis is for
internal user
only

6. REFERENCES

- Aist, G. (1998). Expanding A Time-Sensitive Conversational Architecture For Turn-Taking To Handle Content-Driven Interruption.
- Allen, L., Abella, A., Alonso, T., & Jeremy, H. (2002). Automated natural spoken dialog.
- Avison, D., & Fitzgerald, G. (2003). Information systems development: methodologies, techniques and tools: McGraw-Hill.
- Broadly, C. (2009). Congolese Culture and the French and Kikongo Languages: Congolese Culture and the French and Kikongo Languages: A Linguistic Analysis Jenna Zent EDU 583
- Burgt, S. P. v. d., Andernach, T., Kloosterman, H., Bos, R., & Nijhol, A. (1996). Building Dialogue Systems that Sell (pp. 41-46). New Brunswick: Proceedings Natural Language Processing and Industrial Applications.
- Carlson, R., Edlund, J., Heldner, M., Hjalmarsson, A., House, D., Skantze, G., et al. (2006). Towards human-like behavior in spoken dialog systems.
- Choi, E. (2004). Noise robust front-end for ASR using spectral subtraction, spectral flooring and cumulative distribution mapping.
- Cook, S. (2003). Speech Recognition HOWTO. Home Page.
- Dahl, M., & Claesson, I. (1999). Acoustic noise and echo cancelling with microphone array. Vehicular Technology, IEEE Transactions on, 48(5), 1518-1526.
- Di Fabrizio, G., Tur, G., & Hakkani-Tür, D. (2004). Bootstrapping spoken dialog systems with data reuse.
- Eckert, Penelope. 2000. Linguistic variation as social practice. Oxford: Blackwell. <http://www.stanford.edu/~eckert/csofp.html>
- Fernández, R., Corradini, A., Schlangen, D., & Stede, M. (2007). Towards reducing and managing uncertainty in spoken dialogue systems.

- Hieronymus, J., & Dowding, J. (2004). Clarissa spoken dialogue system for procedure reading and navigation: Citeseer.
- J. Hirschberg, J. Liscombe and J. Venditti,(2003). "Experiments in Emotional Speech," Proceedings of the ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition, Tokyo.<http://www.cs.columbia.edu/~julia/files/cv.pdf>
- Kamm, C. (1995). User interfaces for voice applications. Proceedings of the National Academy of Sciences, 92(22), 10031.
- Kementerian Pelajaran Malaysia (2011), Huraian Sukatan Pelajaran Tahun 2 KSSR
- Kim, L. S. (2006). Masking: Maneuvers of Skilled ESL Speakers in Postcolonial Societies. *English in Southeast Asia: prospects, perspectives, and possibilities*, 191.
- Lee S. K. (2001). A qualitative study of the impact of the English language on the construction of the sociocultural identities of ESL speakers. Unpublished doctoral dissertation, College of Education, University of Houston, USA.
- Lee S. K. (2003). Multiple identities in a multicultural world: A Malaysian perspective. *Journal of Language, Identity and Education* 2(3), 137-158.
- Lee S. K. (2005). What Price English? Identity constructions and identity conflicts in the acquisition of English. In Lee Su Kim, Thang Siew Ming and Kesumawati Abu Bakar. *Language and nationhood: New contexts, new realities*. SoLL's UKM: Bangi.
- Lee S. K. (2006). Masking: Maneuvers of Skilled ESL Speakers in Postcolonial Societies. In Azirah Hashim and Norizah Hassan (Eds.) *English in South East Asia: Prospects, perspectives and possibilities*. K. Lumpur: University of Malaya Press.
- Lee S. K. (2008). Masks of identity: Negotiating multiple identities. Paper presented

at JALT International Conference, Tokyo, Japan.

Lee S. K, Thang Siew Ming & Lee King Siong (2007). Border crossings: Moving

across languages and cultural frameworks. Kuala Lumpur: Pelanduk

Publications.

Lee, S. K., Lee, K. S., & Wong, F. F. (2010). The English Language And Its Impact On Identities Of Multilingual Malaysian Undergraduates.

Levin, E., Pieraccini, R., & Eckert, W. (2000). A stochastic model of human-machine interaction for learning dialog strategies. *Speech and Audio Processing, IEEE Transactions on*, 8(1), 11-23.

Littlefield, J., & Broughton, M. (2005). Dual-Type Automatic Speech Recogniser Designs for Spoken Dialogue Systems.

Musa, N. C., Koo, Y. L., & Azman, H. (2012). Exploring English language learning and teaching in Malaysia. *GEMA: Online Journal of Language Studies*, 12(1), 35-51.

Murugesan, V. (2003). Malaysia promotes excellence in English. *ESL MAGAZINE*.

Natarajan, P., Prasad, R., Suhm, B., & McCarthy, D. (2002). Speech-enabled natural language call routing: BBN Call Director.

Nunamaker Jr, J. F., & Chen, M. (1990). Systems development in information systems research.

Pietquin, O., & Renals, S. (2002). ASR system modeling for automatic evaluation and optimization of dialogue systems.

Plannerer Bernd,(2005) Munich, Germany. *An Introduction to Speech Recognition*.
<http://www.speech-recognition.de/pdf/introSR.pdf>

Raux, A. (2008). Flexible turn-taking for spoken dialogue systems. PhD Thesis, Carnegie Mellon University.

Sukatan Pelajaran Bahasa Inggeris KSSRTahun 2 (2011),

[KPMwww.moe.gov.my/bpk/sp_hsp/bi/kssr/sp_bi_kbsr.pdf](http://www.moe.gov.my/bpk/sp_hsp/bi/kssr/sp_bi_kbsr.pdf)

Tomko, S. (2005). Improving user interaction with spoken dialog systems via shaping.

Vaishnavi, V., & Kuechler, W. (2007). Design research in information systems. *wwwisworld.org*, 22(2), 1-16.

Witt & Young, (1997) Interactive pronunciation training. citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.24.rep