

**GENDER AND ACCENT IDENTIFICATION FOR MALAYSIAN ENGLISH  
USING MFCC AND GAUSSIAN MIXTURE MODEL**

**GOH ENG LYN**

**UNIVERSITI TEKNOLOGI MALAYSIA**

GENDER AND ACCENT IDENTIFICATION FOR MALAYSIAN ENGLISH  
USING MFCC AND GAUSSIAN MIXTURE MODEL

GOH ENG LYN

A project report submitted in partial fulfillment of the  
requirements for the award of the degree of  
Master of Computer Science (Information Security)

Faculty of Computing  
Universiti Teknologi Malaysia

JUNE 2013

Dedicated to my family and my dear sister, Emily.

Thank you for everything.

## **ACKNOWLEDGEMENT**

I would like to dedicate utmost gratitudes towards my supervisor, Dr. Shukor Abd. Razak for patiently guiding me throughout the course of this project.

I would also like to extend heartfelt appreciations towards lecturers who committed to share their invaluable experiences, and knowledge, the inspirations sparked tremendous fighting spirit to strive better.

## ABSTRACT

Speaker and speech variability are a challenge in speaker and speech recognition. In the context of Malaysian English speakers, the variability is highly complex due to sociolinguistic and cultural background. Past researches focused on vowel classification of Malaysian English on a small dataset with limited number of speakers and text corpus. Among speaker specific characteristics, gender is the most prominent features, which is followed by accents. This project approached the issues of speaker and speech variability in Malaysian English by proposing identifier that combined the gender and accent aspects. This method is fulfilled by training the classifier with gender-dependent data and accent-prone text corpus. The gender-accent database collected is comprised of 120 speakers categorized into four gender-accent groups namely Malay Female (CF), Malay Male (MM), Chinese Female (CF) and Chinese Male (CM). MFCC algorithm is used to extract the features, while GMM is the algorithm used to model the identifier. Findings from the test results show that female and Chinese speakers have higher degree of distinctiveness compared to other accent-gender groups. Chinese female is the best recognized accent-gender group, meanwhile Malay Male is the least recognized, due to code-mixing of Malay language and English. Optimum configuration of GMM is also studied across 3 different numbers of Gaussians (12, 24, and 32). It is found that 24 is the most optimal configuration of MFCC-GMM. Meanwhile, it is also known that MFCC-GMM performs better than LPC-KNN on noisy dataset. Overall, the MFCC-GMM identifier scored 99.34% gender identification rate, 67.5% accent identification and 65.83% for accent-gender identification task.

## ABSTRAK

Kepelbagaian ciri-ciri pengucap dan pengucapan adalah satu cabaran dalam pengesahan identiti. Dalam konteks penuturan Bahasa Inggeris di Malaysia, kepelbagaian ciri-ciri amat kompleks disebabkan sosiolinguistik dan budaya yang berbagai. Kajian-kajian lalu menumpukan persoalan atas klasifikasi huruf vokal Bahasa Inggeris dengan menggunakan dataset dengan bilangan pengucap terhad. Antara ciri-ciri khusus pengucap, gender adalah yang paling ketara diikuti oleh aksen. Perdekatan projek ini terhadap isu tersebut adalah sistem pegecam yang mengabungkan pengenalan terhadap kedua-dua aspek gender dan aksen dalam pengesahan pengucap. Ia melibatkan melatih pengelas dengan dataset yang mengandungi ciri gender serta teks corpus yang merangsang aksen dalam pengucapan. Pangkalan data aksen-gender yang telah dikumpulkan terdiri daripada 120 pengucap yang telah dikategorikan kepada empat kumpulan, iaitu Wanita Cina, Lelaki Melayu, Wanita Melayu dan Lelaki Cina. Algoritma MFCC telah digunakan untuk mengekstrakan ciri-ciri daripada rakaman ucapan, manakala GMM adalah algoritma pemodelan pegecam. Hasil-hasil kajian telah menunjukkan bahawa antara kumpulan-kumpulan pengucap, ketepatan pengesahan Wanita Cina adalah paling tinggi, manakala Lelaki kaum Melayu mempunyai ketepatan pengesah yang paling rendah. Ini adalah disebabkan oleh pencampuran kod bahasa Melayu dan Inggeris yang menyebabkan keliruan. Di samping itu, konfigurasi optimum GMM juga telah dikaji dengan jumlah Gaussians yang berbeza (12, 24 dan 32), adalah didapati bahawa 24-Gaussian merupakan konfigurasi yang optimum untuk pegecam MFCC-GMM. Sementara itu, prestasi MFCC-GMM berbanding dengan LPC-KNN juga dikaji menggunakan dataset bising. Secara keseluruhannya, pegecam suara MFCC-GMM mencatat kadar pengenalan sebanyak 99.34%, kadar pengenalan kumpulan aksen yang betul pula adalah 67.5% dan sebanyak 65.83% telah dicatatkan oleh pengenalan kumpulan aksen.

## TABLE OF CONTENTS

<b>CHAPTER</b>	<b>TITLE</b>	<b>PAGE</b>
	<b>DECLARATION</b>	II
	<b>DEDICATION</b>	III
	<b>ACKNOWLEDGEMENT</b>	IV
	<b>ABSTRACT</b>	V
	<b>ABSTRAK</b>	VI
	<b>TABLE OF CONTENTS</b>	VII
	<b>LIST OF TABLES</b>	X
	<b>LIST OF FIGURES</b>	XI
	<b>LIST OF APPENDIX</b>	XII
<b>1</b>	<b>INTRODUCTION</b>	
	1.1 Introduction	1
	1.2 Background of Study	3
	1.3 Problem Statement	5
	1.4 Objectives of Study	5
	1.5 Research Questions	6
	1.6 Scope of Study	6
	1.7 Contribution of Study	7
	1.8 Organization of Report	8

<b>2</b>	<b>LITERATURE REVIEW</b>	
2.1	Introduction	9
2.2	Fundamentals of Speech	9
2.2.1	Speech Production and Articulation	10
2.2.2	Speech Structure	10
2.2.3	Speech Articulation	11
2.2.4	Speech Structure and Articulation in Pattern Modelling	13
2.3	Prosodic Features and Acoustic Parameters	13
2.3.1	Pitch	14
2.3.2	Stress	15
2.4	Speaker and Speech Recognition	16
2.5	Speaker and Speech Variability	17
2.6	Approaches to Gender and Accent Identification	19
2.6.1	Training Corpus	20
2.6.2	Feature Extraction	21
2.6.3	Modelling	23
2.6.4	Accent Modelling	24
2.7	Malaysian English	25
2.7.1	Related Literature on Malaysian English Accent Identification	28
2.8	Summary	29
<b>3</b>	<b>RESEARCH METHODOLOGY</b>	
3.1	Introduction	30
3.2	Research Framework	30
3.3	Research Design and Research Activities	33
3.4	Summary	35



<b>4</b>	<b>EXPERIMENTAL DESIGN</b>	
4.1	Introduction	36
4.2	Gender-Accent Identification Framework	36
4.2.1	Conceptual Design	37
4.2.2	MATLAB-based Accent-Gender Identifier	38
4.3	Training Database	42
4.3.1	Selection of Text	42
4.3.2	Data Preparation	43
4.4	Summary	44
<b>5</b>	<b>RESULTS AND DISCUSSION</b>	
5.1	Introduction	45
5.2	Performance Evaluation	46
5.2.1	Experiment 1	46
5.2.2	Experiment 2	51
5.2.3	Experiment 3	53
5.3	Summary	54
<b>6</b>	<b>CONCLUSION</b>	
6.1	Introduction	55
6.2	Conclusions	55
6.3	Future Work	57
	<b>REFERENCES</b>	58
	<b>APPENDIX A</b>	64
	<b>APPENDIX B</b>	66

**LIST OF TABLES**

<b>TABLE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
3.1	Research Activities (Phase 2 and 3)	33
4.1	Statistics of Training and Testing Data Set	44
5.1	Average Mean Score (Open Set and Closed Set)	46
5.2	Confusion Matrix of Accent-Gender Identification	47
5.3	Confusion Matrix of Accent Identification	48
5.4	Confusion Matrix of Gender Identification	49
5.5	Overall Performance of Baseline Identifier	49
5.6	Identification Rate of Identifier with different number of Gaussians	52
5.7	Performance Comparison of MFCC-GMM and LPC-KNN	53

**LIST OF FIGURES**

<b>FIGURE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	Block diagram of MFCC algorithm	22
3.1	Research Framework	31
4.1	Block diagram of Gender Accent Identifier	38
4.2	Pseudo-code of Identifier	39
5.1	Identification Rate for Different Number of Gaussians	52

**LIST OF APPENDIX**

<b>NO.</b>	<b>TITLE</b>	<b>PAGE</b>
A	Research Design	64
B	Comma Gets a Cure	65

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Speech is a vital medium for human interaction. It is natural ability for humans to detect and discern speakers and rich information contained in speech signals. However, to attain similar task of recognizing speakers and speech, machines must emulate human's auditory and perceptive capability, which is a complex tasks that encompasses multiple disciplines including digital signal processing, pattern recognition and linguistics. This project explores a gender and accent identifier that is able to distinguish between different ethnic groups speaking Malaysian English.

A shared problem between speaker recognition and automatic speech recognition (ASR) is the analysis and modeling of speaker variability [1]. Speaker information one can derive from speech signals in general fall into two groups: linguistic (accent, dialect, language and text contents) and speaker-specific or paralinguistic (gender, age, emotion and identity). This information allows speaker profiling, which can be applied in many fields encompassing voice biometrics for physical security, forensic and Interactive voice response (IVR) systems.

The complexity lies in the huge variability of speech characteristics, which can vary immensely across speakers set due to idiosyncrasies [2]. Gender [3-5] and accent [6-9] are among the most substantial speaker variables [10]. The fusion of these two variables also yielded in increased complexity of speech variations as concluded by Martland, *et al.* [11]. Dissimilar speech characteristics between male and female are caused by physiological and acoustical dissimilarities. Male and female have different thickness of vocal fold and length of vocal tract [4, 12], which lead to formant frequencies and fundamental frequency (pitch) differences. On average, female exhibit higher formant frequencies and pitch than male, which has the lowest frequency among male, female and children [5].

Accent variation is evident and measurable in acoustic phonetics [10, 13, 14]. Variation occurred due to many interacting factors such as second language acquisition, geographical background, ethnicity and socio-economic class [8, 10, 12, 15]. As speaker's backgrounds influenced the formation of accent in speech, accent becomes a major indication that characterizes a speaker belonging to a group or at large, a community. Such is the case of Malaysian English, which is a standard non-native variation of English, introduced during the British colonial in Malaysia [16, 17]. Development of Malaysian English as a localized English standard is largely influenced by other local languages namely Malay, Chinese and Indian [13, 14, 18-20]. Among the three, Malay, the national language of Malaysia, is noted to be the most prominent cause of distinctive variations of Malaysian English [13], meanwhile Chinese is the most significant group of Malaysian English speaker due to its large number of speaker [21].

Performance of speaker-independent and text-independent recognition system deteriorates when speaker possesses features that the system is not trained with [6, 7, 22, 23]. This is especially prominent among non-native speakers as current commercial recognition systems are homogenous and poorly accommodate speaker variability on weighty consideration [9, 22]. Effectiveness of speaker and

speech recognition can be enhanced through deploying classifiers that identify speaker's gender and accent information preceding to the recognition process [7, 10]. Preempt knowledge of the speaker enables more reliable selection of acoustic and lexical model [8, 12]. However, speaker information extraction and classification remains a challenging task due to complexity of the domain [1-3, 10]. Speech corpus used for training, features extraction algorithm and statistical modeling algorithm are all crucial factors to baseline performance of the recognition system.

## 1.2 Background of Study

Compared to other variation of English, Malaysian English has not received much attention in studies of speech processing. Speech modeling of languages in Malaysia is only beginning to gain attention in past ten years [24].

Formant frequencies (specifically the first three frequencies: F0, F1, and F2), energy (expanded energy to produces a stressed syllable), and fundamental frequency (also known as pitch) are three prominent speech characteristics investigated and adapted in current proposed and deployed speaker and speech recognition systems [6, 25]. Formant frequencies and pitch are two evident features that differentiate male and female voicing on interacting physiological and phonetic factors [4, 5, 12]. While energy is most notable in differentiating speaking styles inter-gender and inter-language [25, 26]. Energy is the stress of syllables by applying more muscular energy, which is an important indication of differentiated speaking styles of a language from parent norms [25, 27]. The accent identification system proposed by Yunista *et al.* recorded 94.2% as its highest identification rate using a combination of formant analysis and log energy vector. Researchers have investigated the combination of features to achieve as high as possible the robustness of the

recognition system, as seen in Liu and Fung [25], Li, *et al.* [26] and Metze, *et al.* [28]. Martland, *et al.* [11] in comparing ten vowels sound across gender and accent, concluded that accent factors increased the variability of phonological changes.

In the case of Malaysian English, a majority of speakers learn the language in formal educations after already acquiring a first language, usually associated to their ethnicity [18]. Differentiation in pronunciation is eminently dictated by prosodic and phonological deviation during language learning. On a phonemic level, phonological coding becomes inconsistent from native standards due to unfamiliarity with phonological change, hence phonemes substitution occurred [10]. Hence classifying vowels are substantive to Malaysian English accent identification. Vowel classification is utilized by Paulraj, *et al.* [20] in adapting Malaysian English pronunciation to ASR. Additionally, Mohd Yusof and Yaacob [24] also proposed a vowel classification system based on first three formant frequencies (F0, F1 and F2), which are spectrum peaks resulting from resonant frequencies. On the other hand, Yunista *et al.* [9] also use formant frequencies to identify accent.

Mel Frequency Cepstrum Coefficients (MFCC) and Linear Predictive Coefficients (LPC) are two algorithms dominantly used in speech analysis. LPC is used in all three Malaysian English accent classification systems reported [9, 20, 24]. It is also applied in Teixeira, *et al.* [7]'s accent identification system encompassing six European English accents. A common point among these systems is small database (less than 100 speakers). Meanwhile, MFCC are applied on investigation with relatively bigger databases (up to or more than 1000 speakers) [1, 28]. Training and decision making methods adopted are different among proposed Malaysian English accent classification systems, with Paulraj, *et al.* [20] and Mohd Yusof and Yaacob [24] using neural network while Yusnita, *et al.* [9] adapted a classifier based on K-Nearest Neighbor (KNN). Both Li, *et al.* [26] and Metze, *et al.* [28] investigated on multiple algorithms and proposed fusions of algorithms. Li, *et al.* [23] proposed a MFCC and Gaussian Mixture Model (GMM) on preliminary levels



and decision making based on Support Vector Machine (SVM) are deployed later. Meanwhile, Metze, *et al.* [26] proposed a parallel phone recognizer using MFCC, LPC, Linear Discriminant Analysis (LDA), and Bayesian networks. Additionally, Hidden Markov Model (HMM) is also a popular training model, alternative to GMM and neural network, employed by Liu and Fung [25] and Kumpf and King [29].

### **1.3 Problem Statement**

Malaysian English, a formal second language to the Malaysian community is only beginning to gain attention in local research community. To date, researches of Malaysian English focused on accent identification without considering the gender factor. However, a gender-dependent accent identification system is unreported. Therefore, it is meaningful to extend previously proposed accent identification system to adapt gender and accent, the two substantive speaker variables, and establish baseline recognizer for gender-dependent Malaysian English accents identification.

### **1.4 Objectives of Study**

The goal of this project is to establish and evaluate gender and accent identification technique for Malaysian English using Gaussian Mixture Model (GMM) and Mel Frequency Cepstrum Coefficients (MFCC).

This goal is supported by the following objectives:

- (i) To identify corpus text suitable to elicit gender and accent features in speech.
- (ii) To establish a gender and accent baseline recognizer.
- (iii) To evaluate the performance of the established baseline recognizer.

## **1.5 Research Questions**

This project addresses the following research questions:

- (i) What are the gender-sensitive and accent-sensitive speech features?
- (ii) Does MFCC and GMM-based algorithm increase the performance of gender and accent identification system?
- (iii) What are the performance of the MFCC and GMM-based gender and accent identifier?

## **1.6 Scope of Study**

Scopes of this project are listed as follow:

- (i) Ethnicity and gender are the two concerned variability in this project.
- (ii) Ethnic group involved are Chinese and Indian. Speech data is collected from both male and female. Background, age and geographical location are not considered.

- (iii) This project studies speaker/speech recognition on the acoustic and phonetics aspects. Phonology attributions are not discussed.
- (iv) Recording texts for both training and testing data are excerpted from “Comma Gets a Cure” passage, which is specially designed for pronunciation study [30, 31].
- (v) Level of literacy is not considered, however speakers must be able to read firmly and clearly to maintain relevancy and intelligibility of the speech recording. To foster smooth recording, speakers are given time to read the script before recording.
- (vi) Speech recording environment and speakers are random.
- (vii) Speech sample is recorded with Sony ICD-AX412F digital recorder; storage of speech sample is in mp3 format. Noise-cut function on device is enabled.
- (viii) Pre-processing of audio file includes format, sampling rate and channel conversion. Manual pre-processing is performed using Sony Sound Forge 9.0.

## **1.7 Contribution of Study**

Researches on speaker and speech recognition are majorly focused on the British and American English variations, while relevant studies involving Malaysian English are limited. Findings on Malaysian English accent and speaker gender identification can be extended to application in voice biometric and forensic investigation, where success rate of identification or verification is affected by the diffusion of voice characteristics of speakers set.

## **1.8 Organization of Report**

Chapter 1 of this report introduces the project aim, background of study and problem statement. A detail report of literature review is covered in Chapter 2, followed by discussion on research methodology in Chapter 3. Framework of gender and accent identification is explained in Chapter 4, meanwhile results and findings are discussed in Chapter 5. Conclusion of this study is presented in Chapter 6.

## REFERENCES

- [1] K. Chen and A. Salman, "Learning Speaker-Specific Characteristics with a Deep Neural Architecture," *Neural Networks, IEEE Transactions on*, vol. 22, pp. 1744-1756, 2011.
- [2] S. Y. Lung, "Speaker Recognition," in *Advances in Audio and Speech Signal Processing: Technologies and Applications*, H. Perez-Meana, Ed., ed United States of America: Idea Group Publishing 2007, pp. 371-407.
- [3] C. Huang, T. Chen, S. Li, E. Chang, and J. Zhou, "Analysis of speaker variability," in *Proc. of Eurospeech*, 2001, pp. 1377-1380.
- [4] K. Rakesh, S. Dutta, and K. Shama, "Gender Recognition using speech processing techniques in LABVIEW," *International Journal of Advances in Engineering & Technology*, vol. 1, pp. 51-63, 2011.
- [5] K. Wu and D. G. Childers, "Gender recognition from speech. Part I: Coarse analysis," *The journal of the Acoustical society of America*, vol. 90, p. 1828, 1991.
- [6] C. Huang, E. Chang, and T. Chen, "Accent Issues in Large Vocabulary Continuous Speech Recognition (LVCSR)," Microsoft Technical Report, 2001.
- [7] C. Teixeira, I. Trancoso, and A. Serralheiro, "Accent identification," in *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, 1996, pp. 1784-1787.
- [8] A. Hanani, M. J. Russell, and M. J. Carey, "Human and computer recognition of regional accents and ethnic groups from British English speech," *Computer Speech & Language*, 2012.

- [9] M. Yusnita, M. Paulraj, S. Yaacob, S. A. Bakar, and A. Saidatul, "Malaysian English accents identification using LPC and formant analysis," in *Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on*, 2011, pp. 472-476.
- [10] J. H. L. Hansen and L. M. Arslan, "Foreign accent classification using source generator based prosodic features," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, 1995, pp. 836-839.
- [11] P. Martland, S. P. Whiteside, S. W. Beet, and L. Baghai-Ravary, "Analysis of ten vowel sounds across gender and regional/cultural accent," in *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, 1996, pp. 2231-2234.
- [12] J. C. Thankappan and S. M. Idicula, "Language independent voice-based gender identification system," in *Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India*, 2010, p. 23.
- [13] S. Tan, "Lexical borrowing in Malaysian English: Influences of Malay," ed: Lexis, 2009.
- [14] S. Sulong, "A Closer Look at Malaysian English Vowels: English Meets Malay Dialects."
- [15] M. Huckvale, "ACCDIST: a metric for comparing speakers' accents," 2004.
- [16] N. Ahmad Mahir, S. Jarjis, and M. Kibtiyah, "The use of Malay Malaysian English in Malaysian English: Key considerations," 2007.
- [17] J. Thirusanku and M. M. Yunus, "The Many Faces of Malaysian English," *ISRN Education*, vol. 2012, 2012.
- [18] S. Pillai, "Speaking English the Malaysian way—correct or not?," *English Today*, vol. 24, pp. 42-45, 2008.
- [19] H. San Phoon and M. Maclagan, "The Phonology of Malaysian English: A Preliminary Study," 2009.
- [20] M. Paulraj, Y. Sazali, A. Nazri, and S. Kumar, "A speech recognition system for Malaysian English pronunciation using Neural Network," 2009.

- [21] H. S. Phoon and M. Maclagan, "Chinese Malaysian English Phonology," *Asean Englishes*, vol. 12, pp. 20-45, 2009.
- [22] J. Ortega-Garcia, J. Gonzalez-Rodriguez, and S. Cruz-Llanas, "Speech variability in automatic speaker recognition systems for commercial and forensic purposes," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 15, pp. 27-32, 2000.
- [23] C. Teixeira, I. Trancoso, and A. Serralheiro, "Recognition of non-native accents," in *Proc. Eurospeech*, 1997, pp. 2375-2378.
- [24] S. Mohd Yusof and S. Yaacob, "Classification of Malaysian vowels using formant based features," *Journal of ICT*, vol. 7, pp. 27-40, 2008.
- [25] W. K. Liu and P. Fung, "Fast accent identification and accented speech recognition," in *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, 1999, pp. 221-224.
- [26] M. Li, K. J. Han, and S. Narayanan, "Automatic speaker age and gender recognition using acoustic and prosodic level information fusion," *Computer Speech & Language*, 2012.
- [27] Y. Konig and N. Morgan, "GDNN: A gender-dependent neural network for continuous speech recognition," in *Neural Networks, 1992. IJCNN., International Joint Conference on*, 1992, pp. 332-337.
- [28] F. Metze, J. Ajmera, R. Englert, U. Bub, F. Burkhardt, J. Stegmann, C. Muller, R. Huber, B. Andrassy, and J. G. Bauer, "Comparison of four approaches to age and gender recognition for telephone applications," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 2007, pp. IV-1089-IV-1092.
- [29] K. Kumpf and R. W. King, "Automatic accent classification of foreign accented australian english speech," in *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, 1996, pp. 1740-1743.
- [30] D. Honorof, J. McCullough, and B. Somerville. (2000, 1st December). *Comma gets a cure: A diagnostic passage for accent study*. Available: <http://web.ku.edu/~idea/readings/comma.htm>

- [31] International Dialects of English Archive, (IDEA). (1st December). *Comma Gets A Cure*. Available: <http://web.ku.edu/~idea/readings/comma.htm>
- [32] L. Rabiner and B. H. Juang, "Fundamentals of speech recognition," 1993.
- [33] I. McLoughlin, *Applied speech and audio processing: with Matlab examples*: Cambridge University Press, 2009.
- [34] K. Lodge, *A critical introduction to phonetics*: Continuum, 2009.
- [35] CCRMA, Stanford University. (1999, 1st December 2012). *Speech Production*. Available: <https://ccrma.stanford.edu/CCRMA/Courses/150/speech.html>
- [36] J. Tepperman and S. Narayanan, "Automatic syllable stress detection using prosodic features for pronunciation evaluation of language learners," in *Proc. ICASSP*, 2005, pp. 937-940.
- [37] P. Rose, "Forensic speaker discrimination with Australian English vowel acoustics," *ICPhS XVI Saarbrücken*, pp. 6-10, 2007.
- [38] D. Loakes, "Front vowels as speaker-specific: Some evidence from Australian English," in *Proceedings of the Australian International Conference on Speech Science*, 2004, pp. 289-294.
- [39] H. N. Ting and J. Yunus, "Speaker-independent Malay vowel recognition of children using multi-layer perceptron," in *TENCON 2004. 2004 IEEE Region 10 Conference*, 2004, pp. 68-71.
- [40] B. G. Evans and P. Iverson, "Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences," *The journal of the Acoustical society of America*, vol. 115, p. 352, 2004.
- [41] Y. Wang and L. Guan, "An investigation of speech-based human emotion recognition," in *Multimedia Signal Processing, 2004 IEEE 6th Workshop on*, 2004, pp. 15-18.
- [42] J. Hou, Y. Liu, T. F. Zheng, J. Olsen, and J. Tian, "Using cepstral and prosodic features for Chinese accent identification," in *Chinese Spoken Language Processing (ISCSLP), 2010 7th International Symposium on*, 2010, pp. 177-181.



- [43] C. Sobin and M. Alpert, "Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy," *Journal of psycholinguistic research*, vol. 28, pp. 347-365, 1999.
- [44] M. V. Chan, X. Feng, J. A. Heinen, and R. J. Niederjohn, "Classification of speech accents with neural networks," in *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference on*, 1994, pp. 4483-4486.
- [45] D. Gerhard, *Pitch extraction and fundamental frequency: History and current techniques*: Regina: Department of Computer Science, University of Regina, 2003.
- [46] M. R. Kim, "The phonetics of stress manifestation: Segmental variation, syllable constituency and rhythm," 2011.
- [47] A. Rosenberg and J. Hirschberg, "On the correlation between energy and pitch accent in read english speech," *Proceedings of InterSpeech2006*, 2006.
- [48] K. Malhotra and A. Khosla, "Automatic identification of gender & accent in spoken Hindi utterances with regional Indian accents," in *Spoken Language Technology Workshop, 2008. SLT 2008. IEEE*, 2008, pp. 309-312.
- [49] R. Jacobson, "Language mixing in multilingual Malaysia," *Newsletter of the Research Council of Sociolinguistics*, 2004.
- [50] A. Stolcke, E. Shriberg, L. Ferrer, S. Kajarekar, K. Sonmez, and G. Tur, "Speech recognition as feature extraction for speaker recognition," in *Signal Processing Applications for Public Security and Forensics, 2007. SAFE'07. IEEE Workshop on*, 2007, pp. 1-5.
- [51] A. K. Mohan, "Combining speech recognition and speaker verification," Rutgers, The State University of New Jersey, 2008.
- [52] S. Kajarekar, N. Malayath, and H. Hermansky, "Analysis of speaker and channel variability in speech," in *Proceedings of the IEEE Workshop on Speech Recognition and Understanding*, 1999.
- [53] M. Sedaaghi, "A comparative study of gender and age classification in speech signals," *Iranian Journal of Electrical & Electronic Engineering*, vol. 5, pp. 1-12, 2009.

- [54] T. Chen, C. Huang, E. Chang, and J. Wang, "Automatic accent identification using Gaussian mixture models," in *Automatic Speech Recognition and Understanding, 2001. ASRU'01. IEEE Workshop on*, 2001, pp. 343-346.
- [55] D. Honorof, "Dude! Your goose is named after a punctuation mark. What gives?," in *Doug Honorof Dialect Couch Weblog* vol. 2012, ed, 2012.
- [56] J. C. Wells, *Accents of English* vol. 1: Cambridge University Press, 1982.
- [57] G. Subramaniam, "The changing tenor of English in multicultural postcolonial Malaysia," *3L: Language, Linguistics and Literature, The Southeast Asian Journal of English Language Studies.*, vol. 13, pp. 53-75, 2007.
- [58] P. Collins and X. Yao, "Aspects of the Verbal System of Malaysian English and Other Englishes."
- [59] D. Sturim, W. Campbell, and D. Reynolds, "Classification methods for speaker recognition," in *Speaker Classification I*, ed: Springer, 2007, pp. 278-297.
- [60] S. Suárez-Guerra and J. Oropeza-Rodriguez, "Introduction to Speech Recognition," *Advances in Audio and Speech Signal Processing: Technologies and Applications*, H Perez-Meana editor, Idea Group Publishing, pp. 325-347, 2007.
- [61] J. Cheshire, *English around the world: Sociolinguistic perspectives*: Cambridge University Press, 1991.
- [62] S. Pillai and U. di Malaya, "Prosodic marking of information structure in Malaysian English," 2012.
- [63] K. Rajandran, "English in Malaysia: Concerns facing Nativization."
- [64] K. Ariffin and M. S. Husin, "Code-switching and Code-mixing of English and Bahasa Malaysia in Content-Based Classrooms: Frequency and Attitudes," *The Linguistics Journal*, vol. 5, pp. 220-247, 2011.
- [65] J. Rajadurai, "The faces and facets of English in Malaysia," *English Today*, vol. 20, pp. 54-58, 2004.
- [66] H. S. Phoon, "The Phonological Development of Malaysian English Speaking Chinese Children: A Normative Study," 2010.

- [67] A. P. R. Bodade, M. Kush, M. S. Dessai, M. R. SM, and M. Y. Singh, "SHARKS Speaker Recognition System," in *FCE MCTE SODE: 93 Final Year Project*, ed: MATLAB Central.
- [68] R. Reppen, "Building a corpus: What are the key considerations?," in *The Routledge Handbook of Corpus Linguistics*, A. O'Keeffe and M. McCarthy, Eds., ed New York: Routledge, 2010.