COMPUTER-BASED STUTTERED SPEECH DETECTION SYSTEM USING HIDDEN MARKOV MODEL

CHIN WEE LIP

UNIVERSITI TEKNOLOGI MALAYSIA

COMPUTER-BASED STUTTERED SPEECH DETECTION SYSTEM USING HIDDEN MARKOV MODEL

CHIN WEE LIP

A thesis submitted in fulfillment of the
requirements for the award of the degree of
Master of Engineering (Biomedical)

Faculty of Health Science and Biomedical Engineering
Universiti Teknologi Malaysia

AUGUST 2012

To my beloved family and friends

ACKNOWLEDGEMENT

First and foremost, i shall begin at the beginning, by thanking my original supervisor, the Dr. Tan Tian Swee. To him I am deeply grateful, for his valuable guidance, advice and suggestion. To me, he is more like my friend than my supervisor that continuously encourage me. Thanks to him for giving me the flexibility in doing the research.

A special thanks to Dr. Etain Vong for her many professional especially providing many valuable advices to my understanding of the concept of stuttering disorder. Besides it, I also very appreciate the guidance and emotion support given by En. Ahmad Kamarul Ariff bin Ibrahim throughout this project. He gave a lot of the advice and technical support to me especially the understandings of the speech recognition technique and Hidden Markov Model during the project development. I am also thankful to my friends, thank you for the given knowledge and co-operation to my research studies.

Lastly, my great appreciation also dedicated to my family who has support me all these years, thank you very much for their love and encouragement which always make me possesses strong determination to achieve any success in my life.

ABSTRACT

Stuttering has attracted extensive research interests over the past decades. Most of the available stuttering diagnostics and assessment technique uses human perceptual judgment to overt stuttered speech characteristics. Conventionally, the stuttering severity is diagnosed by manual counting the number of occurrences of disfluencies of pre-recorded therapist-patient conversation. It is a time-consuming task, subjective, inconsistent and easily prone to error across clinics. Therefore, this thesis proposes a computerized system by deploying HMM-based speech recognition technique to detect the stuttered speech disfluency. The continuous Malay digit string has been used as the training and testing set for fluency detection. Hidden Markov Model (HMM) is a robust and powerful statistical-based acoustic modeling technique. With their efficient training algorithm (Forward-backward, Baum-Welch algorithms) and recognition algorithm, as well as its modeling flexibility in model topology and other knowledge sources, HMM has been successfully applied in solving various tasks. In this thesis, a set of normal voice for digit string as database is used for training HMM. Then, the pseudo stuttering voice was collected as testing set for proposed system. The generated experimental results were compared with the results made by Speech Language Pathologist (SLP) from Clinic of Audiology and Speech Sciences of Universiti Kebangsaan Malaysia (UKM). As a result, the proposed system is proven to be capable to achieve 100% average syllable repetition detection accuracy with 86.605% average sound prolongation detection accuracy. The SLP agreed with the result generated by the software. This system can be further enhanced for detecting stuttering disorder for daily speaking words where Microsoft Visual C++ 6.0 and Goldwave have been used for developing the software which can be executed under the window-based environment.

ABSTRAK

Sejak beberapa dekad yang lalu, kegagapan telah menarik minat bagi para-para pengkaji dan ahli terapi. Kebanyakan ujian diagnostik bagi kes-kes gagap masih menggunakan pemerhatian dan persepsi dari pakar atau ahli patologi untuk mengesan tahap keterukan kegagapan. Biasanya, ujian tersebut dilaksanakan secara konvensional dengan mengira jumlah ketidaklancaran percakapan yang dilakukan oleh penyakit. Teknik ini sangat memakan masa, terlalu subjektif, tidak konsisten dan mudah terjadinya ralat yang disebabkan oleh faktor kepelbagaian klinik. Oleh itu, tesis ini mencadangkan suatu sistem komputer yang menggunakan teknik pengecaman suara berasaskan *Hidden Markov Model (HMM)* untuk mengesan ketidaklancaran pertuturan seseorang pesakit. HMM merupakan sebuah teknik permodelan akustik berasaskan statistik yang popular dan canggih. Dengan mengaplikasikan algoritma-algoritma permodelan seperti *Baum-Welch, Forward-backward* dan serta teknik permodelan topologi serta yang lain-lain., ia telah menyelesaikan pelbagai masalah yang dihadapi dalam projek ini. Sebuah pengkalan data yang mengandungi set suara pertuturan rentetan digit melayu yang berterusan dan normal telah dikumpul bagi latihan HMM. Seterusnya, satu set suara *pseudostuttering* telah dikumpul sebagai pangkalan data suara untuk menguji keberkesanan fungsi sistem yang dicadangkan ini. Keputusan eksperimen dan penyelidikan telah dibandingkan dengan keputusan yang dihasilkan oleh ahli patologi dari Klinik Audiologi dan Sains Pertuturan, Universiti Kebangsaan Malaysia. Sistem cadangan ini telah terbukti mampu mengesan sikap kegagapan secara automatik iaitu mencapai purata ketepatan sebanyak 100% dalam mengesan bunyi pengulangan suku kata dan 86.605% dalam mengesan bunyi pemanjangan. Walaubagaimanapun, sistem cadangan ini masih dapat dipertingkatkan lagi dengan menggunakan vokabulari pertuturan harian di mana Microsoft Visual C++ 6.0 dan Goldwave telah digunakan untuk pembangunan projek ini.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GMMs	-	Gaussian Mixture Models
HMM	-	Hidden Markov Model
SLP	-	Speech Language Pathologist
MFCC	-	Me-Frequency Cepstral Coefficient
GUI	-	Graphical User interface
DCT	-	Discrete Cosine Transform
SP	-	Sound Prolongation
SR	-	Syllable Repetition

LIST OF SYMBOLS

π_i	-	Initial state distribution
λ	-	HMM model
μ_{jm}	-	Mean vector for the mth mixture component at state j
$\delta_t(i)$	-	HMM best score
$\alpha_t(i)$	-	Forward variable
$\beta_t(i)$	-	Backward variable
a_{ij}	-	State transition probability
a_p	-	LPC coefficients
A	-	State transition probability distribution
API	-	Application Programming Interface
$b_j(k)$	-	Emission probability
B	-	Emission probability distribution
C_{jk}	-	Weight coefficient for the mth mixture component at state j
F_0	-	Fundamental frequency
M	-	Number of HMM observations symbols per state
$mel(f)$	-	Mel-scale frequency
O	-	Observation probability
q_t	-	HMM state at time *t*
S	-	HMM state
$W(n)$	-	Hamming window function
E	-	Energy
$b_j(x)$	-	Continuous probability density function (pdf)
N	-	Number of HMM States
Σ_{jm}	-	Covariance Matrices

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of Research

Stuttering is a fluency disorders associating with the disruption or disfluency of speech flow in a person (Yaruss and Reardon, 2004). The patients suffer from this disorder confront with difficulties in communication. Thus, they invariably deal with embarrassment attributable to the anxiety and fear in communicating. There is no explicit cause of stuttering in a person. A variety of causes have been suspected but scientists believe that many forms of stuttering are genetically induced (Guitar, 2006).

There are various kinds of speech disfluency types which include interjections, revisions, incomplete phrases repetitions, word repetitions, part-word-repetitions, prolonged sounds, broken words and so on, however, several types of the disfluency speech are related to stuttering. The sound or syllable repetition, monosyllabic whole word repetition, sound prolongation and blocks are categorized as stuttering primary behaviors (Ambrose, 1999).

In stuttering therapy field, the pathologists usually implement an assessment session by using manual counting of the occurrence frequency of speech disfluencies exhibited by patients; the frequency of speech disfluency is one of the criteria that reflect the recovery progression during speech therapy treatment. Conventionally, the assessment is performed by reviewing the recorded tape while counting the number of occurrences of disfluencies speech during the therapist-patient conversation. However, it is time-consuming, subjective, inconsistent and error-prone across clinics (Cordes et al., 1992; Kully and Boberg, 1988). The influence of a clinician's subjective opinion as each clinic using slightly different definitions of stuttering events, and mistakes in the counts (Yaruss, 1997) lead to the inconsistency in assessment. Nevertheless, stuttering has become an interest subject of researchers and pathologist from various domains such as signal analysis, speech pathologist and so on. More and more therapy programs such as Overall Assessment of the Speaker Experience of Stuttering (OASES), Stuttering Severity Inventory (SSI), Lidcombe Program and other tools have been proposed and studied. Less effort has been done on the research of computer-based stuttering speech diagnostics and assessment system (Ooi, 2007). Thus, in this work project, the studies of the stuttering etiology such as nature and characteristics of stuttering, existing stuttering diagnostics and severity tools were studied in order to deal with the underlying problems by developing an automated Malay stuttered speech detection system which was dedicated to provide the more objective supplementary tools to improve the weakness of the perceptual-based conventional stuttering diagnostics methods.

1.2 Problem Statements

Stuttering has become the interesting research issues over past decades. Professionals from different domains exerted utmost effort in developing stuttering assessment methods for the purpose to optimize the performance of the assessment the stuttering behavior such as Stuttering Severity Instrument (SSI-3), Lidcombe Program and others. However, each of the existing conventional assessment technique has its own availability, validity, and inadequateness; the existing problems are discussed as below:

a) Reliability

Researchers and clinicians have been debating on how to best diagnose the stuttering behaviors based on the aspect of the validity and reliability concerns (Carol, 1998). It includes the attempts to identify and quantify the number of occurrence of observable stuttering behaviors accurately such as involuntary, audible or silent, repetitions or prolongations in the utterance of short speech elements (Wingate, 1964). However, such perceptual-based conventional counting measurement may be reliable within clinics but may not consistent across clinics (Kully and Boberg, 1988). Indeed, Cordes and Ingham (1995) demonstrated relative agreement among professionals who work at the same facility but relative disagreement among professionals who work at the different facilities. In addition, the research examination of Packman *et al.* (1993) discovered that seven clinicians who work in the same clinic found that the number of stutters counted from set of given speech samples differ when the clinicians re-counted the same samples at a later time(Thomas and Howell, 2001), the inconsistency of stuttering severity rating was being. Cordes *et al.* (1992) also reported that the degree of experience of a judge in detecting stuttering speech disfluency forms becomes an important factor to ascertain the accuracy of stuttering severity rate of person who is stutter(PWS).

b) Time-Consuming, Subjective and Error-prone

The conventional method of the assessment is to count and classify the observed stuttering behaviors in pre-recorded sample manually by speech therapist. It includes the analysis of sound of dysfluency types such as fast repetition, short pause, unusual lengthening and so on. However, the manual analytical method is a time-consuming task; it is because therapist needs to vacate excessive time that involves the process of manually counting the number of disfluency occurrences in pre-recorded sample. However, such counting processes are subjective and error-prone. Therefore, an automated robust system is required to provide more objective and accurate results instead of using perceptual-based counting methods; the computer-based is also capable to generate a large quantity of results within a short period of time to enable therapist vacate more time on interaction with patients in such a way to monitor their recovery progression more effectively.

c) Cost

The computer-based stuttering assessment system is capable to provide the cost benefit expectation. It can optimize the use of therapists by reducing the numbers of staff involve in analyzing the speech stuttering behaviors of patient. Generally, the cost of the therapy session is high and becomes a burden for the patients; the longer assessment is taken for a patient, the higher of the therapy cost (Ooi, 2007). There are group of individuals that are not affordable for the long-term therapy process, because some of the insurance companies does not cover the rehabilitation fees especially for the fluency disorders in their insurance plan or policy. Thus, proposed system can help patient to reduce the duration of diagnostic session and as a result to minimize the cost of a clinical session.

1.3 Objectives

The main objective of the thesis is to develop a computer-based Malay stuttered speech detection system with high accuracy that is able to mend the inconsistency problems of the conventional perceptual judgments on occurrences of the stuttering behaviors. It includes speech recognition, identification and classification of various kinds of stuttering speech behaviors such as syllable repetition, sound prolongation and so on with high accuracy.

The second objective of the thesis is to design a phonetically balance Malay pronunciation dictionary for any acoustic recognition units in consideration of the robustness of acoustic modeling.

Third objective of the thesis is to design a language model decoding network (speech decoder) that is feasible to recognize and detect various stuttering speech behaviors.

Finally, an automated malay stuttered speech detection report has been designed carefully for the speech therapist to follow closely the recovery progression of people who is stutter(PWS).

1.4 Scopes

In recent years, the research development of speech recognition has achieved a matured level. However, In Malaysia, the domain of applying the speech recognition technology into Malay stuttering therapy field is still in the preliminary stage. Far less effort has been done on the computer-based stuttering assessment system. If the project is to be started from scratch, it would take a very long time before the malay stuttering recognition and detection system can be finalized, used for detail and in-depth study. Alternatively, utilizing the available development tool give a good jump-start towards the creation of the automated stuttering recognition and detection system, and shorten the development time. Therefore, a hidden markov toolkit (HTK) engine has been selected to be utilized during the development of this master thesis. Due to time and contextual constraints, the scope of the thesis has been dedicated as below:

1. Existing of multiple types of the stuttering speech behaviors can be exhibited by stutterer. However, the study focuses on creating a stuttered speech recognizer which is capable to detect two core stuttering behaviors associated with syllable repetition and sound prolongation with high recognition accuracy.
2. Mel Frequency Cepstrum Coefficient (MFCC) is utilized to extract speech features by converting the waveforms to the parameterize forms for the use of HMM.

3. The stochastic-based continuous density hidden Markov model (CDHMM) with Gaussian mixture is used for acoustic modeling.

4. Embedded Baum-Welch training integrated with forward-backward algorithm is applied for re-estimate and compute the optimal parameters of each HMM.

5. The vocabulary consisting of Bahasa Malaysia continuous utterances which are digit zero ("kosong") to nine ("sembilan") (see Appendix E) is collected in quite office environment with 16 bits format and sampled at 16 kHz using fixed high-quality microphone Shure model SM48.

6. Pseudostuttering database which consists of Malay continuous digit stuttered utterance associated with repetition of sound and sound prolongation is collected and verified by Dr. Etain Vong who is the stuttering specialist from Clinic of Audiology and Speech Sciences in UKM for the purpose to ensure the collected pseudostuttering samples are analogous to the real stuttered phenomenon for research purpose in this project work.

7. All the formulated experiments are based on Malay continuous digit speech domain.

8. Spastic Children Association of Johor and Clinic of Audiology and Speech Sciences , UKM assisted in this work project, as well as giving professional feedback.

1.5 Outlines of the Thesis

This thesis is divided into 6 chapters. The first chapter is the introduction of the thesis which explains the problem background, defines problem statement and outlines the project objectives and scopes. In chapter II, the studies is given on the overview of stuttering disorders, followed by the characteristics of the core behaviors and secondary behaviors of stuttering, assessment methods used by pathologists, descriptions of the available stuttering severity tools and so on are discussed in this chapter.

Chapter III reviews the concepts of the speech recognition technique. Acoustic modeling and classification technique like CDHMM, N-gram language model decoding technique, speech signal processing technique and software tools used in this project will be describes in this chapter. Furthermore, the methodology and the design of the proposed system throughout the thesis are explained in Chapter IV. Chapter V discussed all the results obtained in the experiments. The last chapter, Chapter VI gives conclusion and future works to be done.

REFERENCES

Aarnio, T. (1999). *Speech Recognition with Hidden Markov Models in Visual Communication*, Master Degree, University of Turku.

Ambrose, N. G., Yairi, E. (1995). The Role of Repetition Units in the Differential Diagnosis of Early Childhood Incipient Stuttering. *American Journal of Speech-Language Pathology*. 4, 82-88.

Ahmad K. (2008). *Speaker Adaptation Based On Cross Match Technique*. Master Degree, Universiti Teknologi Malaysia, Skudai.

Alpay Koc (1999). *Acoustic feature Analysis For Robust speech Recognition.* Master Degree, Bilkent University

Amarin Deemagarn, A. K. (2004). Thai Connected Digit Speech Recognition Using Hidden Markov Models. *SPECOM 2004: 9th Conference Speech and Computer*. 20-22.

Ambrose (1997). The genetic basis of persistance and recovery in stuttering. *Journal of Speech Language, and Hearing Research (JSLHR).* 40, 556-566.

Ana Rita S. Valente, L. M. T. J., 2 (2011). Characteristics of Stuttering-like Disfluencies in Portuguese School Age Children. *9th Congress for People who Stutter (ISA) and 2nd Latin American Congress on Stuttering (AAT), Buenos Aires, Argentina*. 1-10.

Andrew, G., and Ingham, R. (1971). Stuttering: Considerations in the evaluation of treatment. *British Journal of Communication Disorders*. 6, 129-138.

Asmah, H. O. (1983). *The Malay people of Malaysia and their languages*. Kuala Lumpur: Dewan Bahasa Dan Pustaka.

Atal B. S. (1974). Automatic Recognition of Speakers from Their Voices. *Proceedings of the IEEE*. 64, 460–475.

Atal. B. S. (1976). Effectiveness of Linear Prediction Characteristics of the Speech Wave for Automatic Speaker Identification. *Journal of the American Statistical Association*. 55, 1304–1312.

Atal, B. S. and Schroeder, M. R. (1967). Predictive Coding of Speech Signals. *IEEE Conference on Speech Communication and Processing*. 360–361.

Bahl, L. R., Brown, P. F., de Souza, P. V., and Mercer, R. L. (1988). Speech Recognition With Continuous Parameter Hidden Markov Models. *In Proceedings of the International Conference on Acoustics, Speech and Signal Processing*. 40-43.

Bahl, L. R., Brown, P. F., de Souza, P. V., Mercer, R. L. (1986). Maximum Mutual Information Estimation of Hidden Markov Model Parameters for Speech Recognition. *IEEE International conference on Acoustic, Speech, and Signal Processing*. 49-52.

Bahl, L. R., Jelinek, F., Mercer, R. (1983). A Maximum Likelihood Approach to Continuous Speech Recognition. *IEEE Trans, On Pattern Analysis and Machine Intelligence*. 5, 179-190.

Baken, R. J., Orlikoff, R. F. (2000). *Clinical Measurement of Speech and Voice.* San Diego: Singular Publishing Group.

Bloodstein, O. (1987). *A handbook on stuttering*. (4th ed.) Chicago: National Easter Seal Society.

Brighton, A. P. (1989). *Phoneme Recognition by Hidden Markov Modeling*. Master Degree, Ohio University.

Carys Thomas, P. H. (2001). Assessing efficacy of stuttering treatments. *Journal of Fluency Disorders.* 26, 311-333.

Chee, L. S., Ai, O. C., Yaacob, S. (2009). Overview of Automatic Stuttering Recognition System. *Proceedings of the International Conference on Man-Machine Systems (ICoMMS)***.** 5B7-1 - 5B7-6.

Childers, G., Wu, K. (1991). Gender recognition from speech. *J. Acoustical Society of America*. 90, 1841–1856.

Comez, M. A. (2003). *Large Vocabulary Continuous Speech Recognition For Turkish Using HTK*. Master of Science, Middle East Technical University.

Conture, E. G. (1990). *Stuttering*. (2nd ed.) Englewood Cliff, N.J. :Prentice-Hall.

Conture, E. G. (2001). *Stuttering: Its Nature, Diagnosis, and Treatment*. Boston: Allyn & Bacon.

Cordes, A. K., Ingham. R. J., Frank, P. and Ingham, J. C. (1992). Time interval analysis of interjudge and intrajudge agreement for stuttering event judgements. *Journal of Speech and Hearing Research*. 35, 483 – 494.

Davis, D. M. (1940). The relation of repetitions in the speech of young children to cetain measures of language maturity and situational factors: Parts 2 and 3. *Journal of Speech Disorders*. 5, 235-246.

Davis, S. and Mermelsten, P. (1980). Comparison of Parametric Representation for Monosyllabic Word Recognition in Continuous Spoken Sentence. *IEEE Trans. Acoustics, Speech and Signal Processing*. 28(4), pp. 357 – 366,

Deller, J. R., Hansen, J. H. L., Proakis, J. G. (2000). *Discrete-time Processing of Speech Signals*. Pisatawa, N. J.: IEEE Press.

Duran, S. (1997). *Keyword Spotting Using Hidden Markov Models*. Master Degree, Bogazici University.

Forgie J. W. and Forgie, C. D. 1959. Results Obtained From a Vowel Recognition Computer Program. *Journal of Acoustic Society of America*. 11, 1480-1489.

Fraser, M. (2007). *Self-therapy for the stutterer*. (10th ed.) Memphes, Tennessee: Stuttering Foundation of America.

Fu, R. (2007). *A Probabilistic Approach to Processing Imperfect Transcription in Real World Speech Recognition*. Master of Science, University of Edinburgh.

Gregory, H. H. (2003). *Stuttering Therapy: Rationale and Procedures*. Boston: Allyn & Bacon.

Guitar, B. (2006). *Stuttering: an integrated approach to its nature and treatment*. West Camden Street Baltimore: Lippincott Williams & Wilkins.

Hermansky, H. (1990). Perceptual Linear Predictive (PLP) Analysis of Speech. *Journal of the Acoustical Society of America*. 87, 1738-1752.

Huang, X., Jack, M. A. (1989). Unified Techniques for Vector Quantization and Hidden Markov Models Using Semi- continuous Models. *In Proceedings of the International Conference on Acoustics, Speech und Signal Processing*. 639-642.

Huang, X., Alleva, F., Hon, H. W., Hwang, M., Lee, K. F., Rosenfeld, R. (1993). The SPHINX-II Speech Recognition System: An Overview. *Computer Speech and Language*. 2, 137-148.

Hubbard, Carol P. (1998). Reliability of Judgments of Stuttering and Disfluency in Young Children's Speech. *Journal of Communication Disorders.* 31, 245.

Hwang, M. Y. (1993). *Sub-phonetic acoustic modeling for speaker-independent continuous speech recognition.* Doctor Philosophy, Carnegie Mellon University.

Jelinek, F. (1976). Continuous Speech Recognition by Statistical Methods. *Proc of the IEEE.* 64(4), 532-556.

Jelinek, F., (1991). The Struggle for Improved Language Models. *Proc. of Eurospeech,* 1037-1040.

Jia, C. W., Yu, S. W. (2002). Chip design of MFCC extraction for speech recognition. *Integration, VLSI journal.* 32, 111-131.

John H. L. Hansen (1993). *Discrete-Time Processing of Speech Signals*. Macmilan, 634-638.

Johnson, W. (1959). *The Onset of Stuttering*. Minneapolis: University of Minnesota Press.

Joseph Agnello, R. B., Hugo Gregory, J. David Williams, Gerald Moses, Fred (2008). *Advice to those who stutter*. (5th ed.) Memphis: Stuttering Foundation of America.

Juang, B. H., Rabiner, L. R. (1991). Hidden Markov Models. *Technometrics*. 33(3), 251-272.

Juang. B. H. (1984). On the Hidden Markov Model and Dynamic Time Warping for Speech Recognition - A Unified View. *AT& T Technical Journal*. 63, 1213-1243.

Juang. B. H. (1985). Maximum Likelihood Estimation for Mixture Multivariate Stochastic Observations of Markov Chains. *AT&T Technical Journal*. 64, 1235-1249.

Juang. B. H. and Rabiner. L. R. (1985). Mixture Autoregressive Hidden Markov Models for Speech Signals. *IEEE Transactions on Acoustics. Speech and Signal Processing*. 33, 1404-1413.

Juang. B. H. and Rabiner. L. R. (1990). The Segmental k-Means Algorithm for Estirnating Parameters of Hidden Markov Models. *IEEE Transactions on Acoustics. Speech and Signal Processing*. 38, 1639-1641.

Juang. B. H., Rabiner, L. R., Levinson, S. E. and Sondhi, M. M. (1985). Recent Developments in the Application of Hidden Markov Models to Speaker-Independent Isolated Word Recognition. *In Proceedings of the International Conference on Acoustics. Speech and Signal Processing*. 9-12.

Juang. B. H., Rabiner, L. R.. and Wilpon. J. G. (1987). On the Use of Bandpass Liftering in Speech Recognition, *IEEE Transactions on Acoustics. Speech and Signal Processing*. 35, 947-954.

Kully, D. & Boberg, E. (1988). An investigation of interclinic agreement in the identification of fluent and stuttered syllables. *Journal of Fluency Disorders.* 13, 309-318.

Lee, K. F. (1986). Incremental Network Generation in Word Recognition. *In Proc IEEE.*

Lee, K. F. (1988). *Large-Vocabulary Speaker-Independent Continuous Speech Recognition: The SPHINX System.* Doctor Philosophy, Carnegie Mellon University.

Lee, C. H.. Juang, B. H., Soong, F. K., and Rabiner, L. R (1989). Word Recognition Using Whole Word and Subword Models. *In Proceedings of the International Conference on Acoustics. Speech and Signal Processing, New York: IEEE*, 683-686.

Lee, K. F., Hon, H. W. (1989). Speaker Independent Phone Recognition Using Hidden Markov Model. *IEEE Transactions on Acoustics Speech and Signal Processing*. 1641-1648.

Lee, K. F., Hon, H. W., Reddy, R. (1990). An Overview of the SPHINX Speech Recognition System. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 38, 35-45.

Lynn, D. W. and Marcia, A. B. (2000), *Speech Recognition, Speech Signal Processing*. The Electrical Engineering Handbook, CRC Press LLC.

Manning, W. H. (2009). *Clinical Decision Making In Fluency Disorders*. (3rd ed.) New York: Delmar.

Markowitz, J. A. (1996). *Using Speech Recognition*. Upper Saddle River, New Jersey: Prentice Hall, Inc.

Markku Ursin (2002). *Triphone Clustering in Finnish Continuous Speech Recognition*. Master Degree, Helsinki University of Technology.

Matthew Nicholas Stuttle (2003). *A Gaussian Mixture Model Spectral Representation for Speech Recognition*. Doctor of Philosophy, University of Cambridge.

Mohamed M. Azmi, H. T. (2008). Comparative Experiments to Evaluate The Use Of Syllables For The Automatic Recognition Of Arabic Spoken Names In Noisy Environment. *Conference Neural Networks & Signal Processing*.

Mohamed M. Azmi, H. T. (2008). Noise Robustness Using Different Acoustic Units. *ICALIP2008*. 1115-1120.

Mohamed M. Azmi, H. T. (2008). Syllable-Based automatic arabic speech recognition in noisy enviroment. *ICALIP*. 1436-1441.

Mondher Frikha, A. B. H. a. M. L. (2010). Hidden Markov models (HMMs) isolated word recognizer with the optimization of acoustical analysis and modeling techniques. *International Journal of the Physical Sciences*. 6(22), 5064-5074.

Mondher Frikha, A. B. H. a. M. L. (2010). Hidden Markov models (HMMs) isolated word recognizer with the optimization of acoustical analysis and modeling techniques. *International Journal of the Physical Sciences.* 6(22), 5064-5074.

Myers, F. L. (1978). Relationship between eight physiological variables and severity of stuttering. *Journal of Fluency Disorders*. 3, 181-191.

Ooi, C. A. (2007). *Computer-based Malay Stuttering Assessment System*, Master Thesis, Universiti Teknologi Malaysia, Skudai.

Pascal Wiggers (2001). *Hidden Markov Models for Automatic Speech Recognition and their Multimodal Application.* Master Degree, Delft University of Technology

Peter Howell (2011). *Recovery from stuttering*. New York: Psychology Press

Peter R. Ramig, Darrell M. Dodge (2009). *The child and adolescent stuttering treatment and activity resource guide*. (2nd ed.) New York: Delmar, Cengage Learning.

Philip, G and E. S. Young (1987). Man-machine Interaction by voice: Developments in Speech Technology. *Journal of Information Science,* 13, 3-14.

Pindzola, R., Jenkins, M., Lokken, K. (1989). Speaking rates of young children. *Language, Speech, and Hearing Services in Schools*. 20, 133-138.

Ravi Kumar K M. (2011). Comparison of Multidimensional MFCC Feature Vectors for Objective Assessment of Stuttered Disfluencies. *Int. J. Advanced Networking and Applications.* 02(05), 854-860.

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE.* 77(2), 257 –286.

Rabiner, L. R., Wilpon, J. G., Soong, F. K. (1989). High performance Connected Digit Recognition Using Hidden Markov Models. *IEEE International conference on Acoustic, Speech, and Signal Processing*.

Rabiner, L. and Juang, B. H. (1993). *Fundamentals of Speech Recognition*. Englewood Cliffs, N.J.: Prentice Hall.

Ravikumar, K. M., Nagaraj, H. C. (2009). An Approach for Objective Assessment of Stuttered Speech Using MFCC Features. *Digital Signal Processing(DSP).* 9(1), 19-24.

Rao, K Vasudeva (2000). *Speaker Independent isolated Digit Voice Recognition Using Discrete Hidden Markov Model*. Master Degree, Indian Institute of Technology, Kanpur.

Reitzes, P. (2006). Pausing: Reducing the Frequency of Stuttering. *Journal of Stuttering, Advocacy & Research*, 1, 64-78

Reardon, N. A., Yaruss, J. S. (2004). *The Source for Stuttering: Ages 7 - 18*. East Moline, IL: LinguiSystems.

Remzi Serdar (2006). *Isolated word recognition from in-ear microphone data using Hidden Markov Models (HMM)*. Master Degree, Naval Postgraduate School.

Riis, S. K. (1998). *Hidden Markov Models and Neural Networks for Speech Recognition, Technical.* Doctor of Philosophy, University of Denmark.

Riley, G. (1994). *Stuttering Severity Instrument for Children and Adults*. (3rd ed.) Austin, TX: Pro-Ed.

Riper, C. (1982). *The nature of stuttering.* (2 ed.) Englewood Cliffs, NJ: Prentice-Hall.

Rodman, R. D (1999). *Computer Speech Technology*. Norwood: Artech House, Inc.

Seifu, Z. (2003). *Hidden Markov Model Based Large Vocabulary, Speaker Independent, Continuous Amharic Speech Recognition*. Master of Science, Addis Ababa University.

Shapiro, D. A. (1999). *Stuttering Intervention: A Collaborative Journey to Fluency Freedom*. Austin, TX: Pro-Ed.

Shelley B. Brundage, A. K. B., Amy N. Lengeling, Jeffrey J. Evans (2006). Comparing judgments of stuttering made by students, clinicians, and highly experienced judges. *Journal of Fluency Disorders*, 31, 271-283. .

S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, V. Valtchev, P. Woodland(2006). *The HTK Book (for HTK Version 3.4)*. Cambridge University Engineering Department.

Sim, K. C. (2006). *Structured Precision Matrix Modeling for Speech Recognition.* Doctor of Philosophy, University of Cambridge.

Tatavarty, U. R. (2011). *Implementation of numerically stable hidden Markov model*. Master Degree, University of Nevada.

Thangarajan, R. (2008). Syllable-Based Continuous Speech Recognition for Tamil. *South Asian Lnaguage*. 18(1), 71-85.

Thomas, C., Howell, P. (2001). Assessing efficacy of stuttering treatments. *Journal of Fluency Disorders*. 26, 311-333.

Ting, C. M. (2007*). Malay Continuous Speech Recognition Using Continuous Density Hidden Markov Model*. Master Degree, Universiti Teknologi Malaysia.

Tomi Kinnunen (2003). *Spectral Features for Automatic Text-Independent Speaker Recognition*. PHD Thesis, University of Joensuu.

Tunali, V. (2005). *A Speaker Dependent, Large Vocabulary, Isolated Word Speech Recognition System for Turkish*, Master Thesis., Marmara University.

Van Riper, C. (1971). *The Nature of Stuttering*. Englewood Cliffs, N.J.: Prentice-Hall.

Van Riper, C. (1982). *The Nature of Stuttering*. (ed 2) Englewood Cliffs, N.J.: Prentice-Hall.

Ward (2006). *Stuttering and cluttering*. Madison Avenue, New York: Psychology Press.

Wexler, K. B., Mysack, E. D. (1982). Disfluency characteristics of 2-, 4- and 6 year old males. *Journal of Communication Disorders*, 12, 133-145.

William, D. E. (1978). *The problem of stuttering. In F. Darley, and D. Spriestersbach, Diagnostic Methods in Speech Pathology*. New York: Harper & Row.

Wingate, M. E. (1964). A Standard Definition Of Stuttering. *Journal of Speech and Hearing Disorders.* 29, 484-489.

Yairi, E., Ambrose, N. G. (1999). Early Childhood Stuttering I: Persistency and Recovery Rates. *Journal of Speech, Language and Hearing Research*. 42, 1097-1112.

Yairi, E. (1981). Disfluencies of normally speaking two-year old children. *Journal of Speech and Hearing Research*. 24, 490-495.

Yairi, E. (1982). Longitudinal studies of disfluncies in two year old children. *Journal of Speech and Hearing Research*. 25, 155-160.

Yairi, E. (1997a). *Early stuttering*. (2^{nd} ed), Boston: Allyn & Bacon.

Yairi, E. and Ambrose, N. G. (2005). *Early childhood stuttering: For Clinicians by Clinicians*. Austin, TX: Pro-Ed.

Yairi, E., and Lewis, B. (1984). Disfluencies at the oneset of stuttering. *Journal of Speech and Hearing Research*, 27, 154-159.

Yapanel, U. (2000). *Garbage Modeling Techniques For a Turkish Keyword Spotting System*. Master Degree, Istanbu Teknik University.

Yaruss, J.S. (1997). Clinical Measurement of Stuttering Behaviors. *Contemporary Issues in Communication Science and Disorders*. 24, 33-44.

Yaruss (1998). Real-Time Analysis of Speech Fluency. *American Journal of Speech-Language Pathology*. 7.

Yaruss, J. S., Quesal, R. W. (2006). Overall Assessment of the Speaker's Experience of Stuttering (OASES): Documenting multiple outcomes in stuttering treatment. *Journal of Fluency Disorders*, 31, 90-115.

Young (1961). Predicting ratings of severity of stuttering. *Journal of Speech and Hearing Disorders*, Monograph Supplement, 7, 31-54.

Young, S., Evermann, G., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V. and Woodland, P. (2002). T*he HTK Book (for HTK Version 3.2).* Microsoft Corporation and Cambridge University Engineering Department, England.

Yuk, D. (1999). *Robust Speech Recognition Using Neural Network And Hidden Markov Models - Adaptations Using Non-Linear Transformations The State.* Doctor of Philosophy, University of New Jersey.

Zebrowski, P. (1991). Duration of the speech disfluencies of beginning stutterers. *Journal of Speech and Hearing Research.* 34, 483-491.

Zhao, J. (2000). *Network and N-gram Decoding In Speech Recognition*. Master Degree, Mississippi State University.