VIRTUAL SPEECH THERAPY SYSTEM: SPEECH INTERVENTION APPLICATION FOR CHILDREN WITH AUTISM SPECTRUM DISORDER

MARLYN MASERI

FACULTY OF ENGINEERING
UNIVERSITI MALAYSIA SABAH
2023

VIRTUAL SPEECH THERAPY SYSTEM: SPEECH INTERVENTION APPLICATION FOR CHILDREN WITH AUTISM SPECTRUM DISORDER

MARLYN MASERI

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING

FACULTY OF ENGINEERING
UNIVERSITI MALAYSIA SABAH
2023

UNIVERSITI MALAYSIA SABAH
BORANG PENGESAHAN STATUS TESIS

JUDUL : **VIRTUAL SPEECH THERAPY SYSTEM: SPEECH INTERVENTION APPLICATION FOR CHILDREN WITH AUTISM SPECTRUM DISORDER**

IJAZAH : **SARJANA KEJURUTERAAN**

BIDANG : **KEJURUTERAAN KOMPUTER**

Saya **MARLYN MASERI**, Sesi **2016-2023**, mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

☐	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)
☐	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
☑	TIDAK TERHAD	

Disahkan Oleh,

Marlyn

MARLYN MASERI
MK1621031T

ANITA BINTI ARSAD
PUSTAKAWAN KANAN
UNIVERSITI MALAYSIA SABAH

(Tandatangan Pustakawan)

Tarikh : 19 September 2023

(Dr. Mazlina Binti Mamat)
Penyelia

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, equations, summaries, and references, which have been duly acknowledged.

11 July 2023

Marlyn

MARLYN MASERI
MK1621031T

CERTIFICATION

NAME	:	**MARLYN MASERI**
MATRIC NO.	:	**MK1621031T**
TITTLE	:	**VIRTUAL SPEECH THERAPY SYSTEM: SPEECH INTERVENTION APPLICATION FOR CHILDREN WITH AUTISM SPECTRUM DISORDER**
DEGREE	:	**MASTER OF ENGINEERING**
FIELD	:	**COMPUTER ENGINEERING**
VIVA DATE	:	**11 JULY 2023**

CERTIFIED BY;

Signature

SUPERVISOR
Dr. Mazlina Binti Mamat

ACKNOWLEDGEMENT

I have received a great deal of support and assistance throughout this research and thesis writing.

I would first like to thank my supervisor, Dr Mazlina Binti Mamat for her valuable guidance and assistance in completing this research. Your encouragement and patient support pushed me to do it better, further my research and finally completed this thesis. Without your support, it would be impossible for me to complete my research.

Next, I would like to thank my parents and sisters for their encouragement and support all through my studies. I know you are always there for me.

Last but not least, I could not have completed this research without the support of my husband, Hamka. Thank you for your unwavering support, encouragement, and understanding throughout the journey of completing this thesis.

Marlyn Maseri
11 July 2023

ABSTRACT

With technological advancement, the integration of digital technology into education has become possible. Nowadays, parents are more open to adopting and using technology like software applications ("apps") to enrich teaching and learning processes for the benefit of their children. Based on research, there is an increase in the use of apps in early education. Apps are also being utilized as intervention tools for children with Autism Spectrum Disorder (ASD). Most apps are solely to serve as education resources rather than a virtual therapy system, lacking the evidence-based standard teaching: the Applied Behavioural Analysis (ABA). Hence, this research aims to develop a Virtual Speech Therapy System (VSTS), incorporating the ABA strategy, particularly Discrete Trial Training (DTT), for children with ASD. Two key objectives were set for the VSTS to be realized: (1) To transform and implement the conventional DTT into a software application, and (2) To develop a speech recognizer implementing the Mel-frequency Cepstrum Coefficients (MFCC) and Hidden Markov Model (HMM). In achieving the first objective, the practicality of the coding method and design of the DTT-based lesson module is assessed and investigated. Next, a speech recognizer is developed, trained, and tested using 2904 speech samples. The speech recognizer is developed using the Hidden Markov Model (HMM) that utilizes 39 Mel Frequency Cepstral Coefficient (MFCC) features. The speech recognizer achieves 95.2% and 95.0% training and testing accuracies for English speeches, respectively. For Malay speeches, the recognition performance is 94.38% for training and 93.75% for testing, respectively. This research contributes to developing a speech intervention application that employs the ABA technique, namely the DTT. This intervention application is not intended to replace therapists but to enhance learning outcomes. The app, most importantly, can be used as a home-based intervention tool.

ABSTRAK

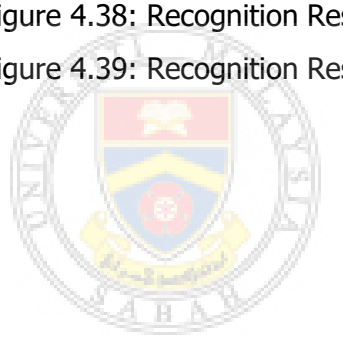
SISTEM TERAPI PERTUTURAN MAYA:
APLIKASI INTERVENSI PERTUTURAN UNTUK KANAK-KANAK DENGAN GANGGUAN SPEKTRUM AUTISMA

Dengan kemajuan teknologi, integrasi teknologi digital ke dalam bidang pendidikan telah menjadi suatu kemungkinan. Pada masa kini, ibu bapa lebih terbuka untuk menerima pakai dan menggunakan teknologi seperti perisian aplikasi ("aplikasi") untuk memperkayakan proses pengajaran dan pembelajaran demi manfaat anak-anak mereka. Berdasarkan kajian, terdapat peningkatan dalam penggunaan perisian aplikasi dalam pendidikan awal. Perisian aplikasi juga boleh digunakan sebagai alat intervensi untuk kanak-kanak yang mengalami Kecelaruan Spektrum Autisma (ASD). Namun demikian, sebahagian besar perisian aplikasi tersebut dicipta untuk berfungsi sebagai medium pengajaran satu hala dan tidak berupaya untuk berfungsi sebagai sistem terapi maya yang melibatkan interaksi dua hala. Selain itu, perisian aplikasi tersebut dibangunkan tanpa menggunapakai konsep pengajaran piawai berasaskan bukti, yang dikenali sebagai Analisis Tingkah Laku Gunaan (ABA). Justeru, penyelidikan ini bertujuan untuk membangunkan Sistem Terapi Pertuturan Maya (VSTS), yang menggabungkan strategi ABA, khususnya Latihan Percubaan Diskret (DTT), untuk kanak-kanak dengan ASD. Dua objektif utama telah ditetapkan untuk VSTS direalisasikan: (1) Untuk mengubah dan melaksanakan DTT konvensional menjadi aplikasi perisian, dan (2) Untuk membangunkan pengecam pertuturan menggunakan gabungan Mel-frequency Cepstrum Coefficients (MFCC) dan Model Markov Tersembunyi (HMM). Dalam mencapai objektif pertama, prosedur DTT telah diubah untuk disesuaikan untuk memenuhi rekabentuk modul pelajaran dalam bentuk perisian aplikasi. Seterusnya, pengecam pertuturan telah dibangunkan. Pengecam pertuturan telah dilatih dan diuji menggunakan 2904 sampel pertuturan dalam Bahasa Inggeris dan Melayu. Pengecam pertuturan dibangunkan menggunakan Model Markov Tersembunyi (HMM) berdasarkan 39 pekali Mel Frequency Cepstral Coefficient (MFCC). Pengecam pertuturan masing-masing mencapai 95.2% dan 95.0% ketepatan latihan dan ujian untuk pertuturan bahasa Inggeris. Bagi Bahasa Melayu, prestasi pengecaman masing-masing adalah 94.38% untuk latihan dan 93.75% untuk ujian. Penyelidikan ini menyumbang kepada pembangunan aplikasi intervensi pertuturan yang menggunakan teknik ABA, iaitu DTT. Aplikasi intervensi ini bukan bertujuan untuk menggantikan ahli terapi tetapi untuk meningkatkan kualiti hasil pembelajaran. Aplikasi ini, yang paling penting, boleh digunakan sebagai alat intervensi pertuturan di rumah.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES



LIST OF SYMBOLS

$\propto$	-	filter coefficient, 0.975
$\boldsymbol{f}$	-	frequency
$\boldsymbol{d_t}$	-	delta coefficients
$\boldsymbol{m_j}$	-	log filter bank amplitudes
μ	-	mean vector
Σ	-	covariance matrix

LIST OF ABBREVIATIONS

ASD	-	Autism Spectrum Disorder
DTT	-	Discrete Trail Training
ABA	-	Applied Behaviour Analysis
EIBI	-	Early Intensive Behavioural Intervention
PRT	-	Pivotal Response Treatment
ESDM	-	Early Start Denver Model
DFT	-	Discrete Fourier Transform
DCT	-	Discrete Cosine Transformation
MFCC	-	Mel-Frequency Cepstral Coefficients
LPC	-	Linear Predictive Coding
LPCC	-	Linear prediction cepstral coefficient
PLP	-	Perceptual Linear Prediction
HMM	-	Hidden Markov Model

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

In 1948, Leo Kanner classified autism as a developmental disorder. Since then, other researchers have taken an interest in autism, which has resulted in ongoing revisions to the autism diagnostic standard (Corsello, 2005). The American Psychiatric Association's fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) was published in May 2013 and includes significant revisions to the diagnostic criteria for autism. The revised criteria include two impairment domains: [A] restricted interests and repetitive behaviours and [B] restricted interests and social interaction. Furthermore, the term "Autism Spectrum Disorder, ASD" refers to all three disorders: autistic disorder, Asperger's disorder, and pervasive developmental disorder - unless otherwise specified (PDD-NOS) (American Psychiatric Association, 2013).

Symptoms of autism spectrum disorder, or ASD, typically appear during the early stages of human development. They struggle with social communication skills such as starting conversations, responding to others' communicative bids, and engaging in reciprocal conversations (Eigsti *et al.*, 2011). According to researchers, approximately 25% to 61% of these children have little or no functional speech (Schlosser & Wendt, 2008), and approximately 25% to 50% of them have not developed language by the time they are 10 to 13 years old (Kasari *et al.*, 2008). Because communication is required in everyday life, learning basic communication skills is essential.

Communication is required to exchange messages, thoughts, feelings, and information. Communication includes speech, vocalisations such as sounds and shouts, body language such as facial expressions and posture, sign language or picture exchange, the use of communication devices, and writing (Vellonen *et al.*, 2012). Over the years, researchers and clinicians noted that Applied Behaviour Analysis (ABA) is an effective treatment for ASD through thousands of published research studies (Handleman & Harris, 2005).

Furthermore, the American Psychological Association (APA) and the US Surgeon General both regarded ABA as an evidence-based "best" practise treatment. According to the MADSEC Autism Task Force Report, ABA is the systematic implementation of interventions based on learning theory principles in order to improve socially significant behaviours to a meaningful degree and to demonstrate that the interventions used are responsible for the behaviour improvement. Socially significant behaviours include reading, academics, social skills, communication, and adaptive living skills (MADSEC Autism Task Force, 1999). The process of ABA therapy requires intensive training by the therapist and supervision by a board-certified behaviour analyst to monitor the children's individualized program. In Malaysia, there are few qualified therapists available in special schools or hospitals provided by the government, which leads to long waiting lists, and forces parents to seek the assistance of private therapists, which are more expensive (Ilias *et al.*, 2017).

Usually, the intervention costs around RM700 to RM1800 per month in a privately-run centre. At the same time, a non-governmental organisation such as the National Autism Society of Malaysia (NASOM) charges a monthly fee of RM350 for a half-day class held five times a week (Sani *et al.*, 2011). In addition, the therapy program is only available in urban areas. While for underprivileged parents who live in rural areas, their hope of getting effective therapy for their children is hard to be fulfilled. In this modern era, new technology frontier has been utilized as a bridge to instil effective communication in children with ASD. Many researchers have ventured into implementing software applications with high-tech devices (e.g., iPad, tablet) as learning interventions. Researches were convinced that future technology could benefit the diagnosis and intervention process of individuals with ASD to improve their quality of life (Bölte *et al.*, 2010).

Several previous review papers have discussed the efficacy of using technology-based interventions to improve communication skills in people with ASD (Boyd *et al.*, 2015; Kagohara *et al.*, 2013; Papathomas & Goldschmidt, 2017; Still *et al.*, 2014; Wainer & Ingersoll, 2011; Wieckowski & White, 2017). Digennaro *et al.*(2011) focused on studies that used technology to help children with ASD learn social skills, and their findings support the notion that technology can help children with ASD learn social skills. The same year, Ramdoss *et al.* (2011) examined the use of computer-based interventions (CBI) to teach communication skills to children with ASD. According to the findings, integrating technology into intervention is a promising practise for improving both vocal and nonverbal communication. Some recent research found that high-tech tools, including smartphone technology, might be employed in educational settings and could be a promising approach as a direct intervention for children with ASD to teach language, emotion recognition, or social skills (Boyd *et al.*, 2015; Kagohara *et al.*, 2013; Papathomas & Goldschmidt, 2017; Still *et al.*, 2014; Wainer & Ingersoll, 2011; Wieckowski & White, 2017).

1.2 Problem statement

The technology-based intervention has become an emerging alternative in providing cheaper and more effective therapy, especially for children living farther from special needs schools or centres. Recently, several software applications (apps) of assistive technology have been developed to improve development skills needed for the ASD children like cognitive skills, communication skills, and others (Achmadi *et al.*, 2012; Kagohara *et al., 2013*; Parsons, 2015; Wojciechowski & Al-Musawi, 2016; Xin & Leonard, 2014). Most of the developed apps could be obtained for free and used on mobile devices like iPad and tablets. Since most parents own mobile devices, intervention using apps could be considered cheaper. Some apps could be obtained for free compared to traditional materials like books, picture-based symbols, and speech-generating devices (SGDs). Researchers have also noted that technology materials like mobile devices has been proven as viable intervention tools, especially

in improving communication skills (Heath *et al.*, 2021; Maseri *et al.*, 2021; Saul & Norbury, 2020; Vandermeer *et al.*, 2013).

Despite numerous learning apps being developed over the years, only a few adopt the ABA frameworks in their system (Gallardo *et al.*, 2021; Grossard *et al.*, 2017; Lian & Sunar, 2021; Rezae *et al.*, 2020). Almost all apps are designed and developed to function as teaching materials, thus unsuitable for children with ASD. For example, HER (Headsprout Early Reading), a tablet-based reading program, teaches oral reading skills to preschool children. Other apps like Proloquo2Go, My Choice Board, Pick A Word, and Go Talk Now were developed to replace SGDs. Therefore, the evidence-based traditional teaching protocol by the teachers must be included in the intervention apps. Having an app that could act as a teacher and provide intervention for ASD children would give a great advantage because it would allow the intervention to be extended outside of the training institution or centre. The app allows for intervention to take place at any time and from any location, and also, reducing training expenditures. Furthermore, early intervention could be extended at home, promoting greater learning as the learning process is not limited to school.

However, developing a virtual therapy system specifically for children with ASD is not easy since the way they learn and understand new things is different compared to the typically developing (TD) children. To improve their skills, they need proper teaching, such as evidence-based lesson procedures like ABA. Hence, integration of technology with the ABA lesson module as a virtual therapy system would bring great advantages in providing intervention to them. When designing technology for children with disabilities, many criteria need to be considered. How they react to the technology, how they could learn from it, whether they can focus during the learning process, and whether they need additional help from other people to use the technology should be considered. The development of a virtual therapy system is not meant to replace the human therapist but to complement it.

1.3 Research Objectives

The goal of this study is to develop a Virtual Speech Therapy System (VSTS) for children with ASD. The VSTS incorporates one of the most effective ABA methods: Discrete Trial Training (DTT) as its learning approach. The VSTS comes with two distinctive features, customizable and simple to operate. To develop the VSTS, this research embarks on two key objectives as follows:

[1] To deploy the conventional Discrete Trial Training into a software application.
[2] To develop a speech recognizer using the Mel-frequency Cepstrum Coefficients (MFCC) feature extraction method and Hidden Markov Model (HMM).
[3] To evaluate the performance of the developed Virtual Speech Therapy System (VSTS).

The adoption and integration of technology into the traditional intervention process is the study's highlight. With this, the manual DTT process is transformed into an automated process by using an algorithm and other technology, such as the Speech Recognizer, to match human interaction. The benefits of this implementation are reducing costs, available anytime and anywhere, and mobility. Nowadays, more people are adapting and embracing digital transformation in business, learning and other processes. Thus, this research is significant in expanding the digital transformation in healthcare, i.e., therapy.

1.4 Scope of Project

The scope of this research is summarized as follows:

1) Target User: The primary target user for this application is children diagnosed with ASD between four to twelve years old. The other users are their parents,

family members, or caregivers. The application is designed with a simple interface, easy-to-access, easy-to-understand, and user-friendly.

2) Lesson Curriculum: The goal of this study is to create a virtual speech therapy system that can provide early intervention to improve the target user's speech. Therefore, the curriculum is limited to early childhood lesson plans and focuses only on basic language skills like listening, reading, and speaking.

3) Design application: The application is designed in the Android framework since most available tablets or smartphones suitable for children are running on the Android platform.

4) Language and geographical: The application focuses on teaching language skills in Bahasa Melayu and English to Malaysian children diagnosed with ASD. Therefore, the languages included in the VSTS are Bahasa Melayu and English. The Speech Recognition algorithm is designed based on the selected language and character of the target user.

1.5 Thesis Organisation

The first chapter, Chapter 1 covers all of the general information about this project. It also provides an overview of this project, including the objectives, project scope, and problem context. Next, Chapter 2 discusses a brief review of the fundamental characteristic of autism, speech recognition algorithm techniques and applications, as well as the ABA lesson module for children with ASD. Various detailed technology-based interventions for speech are reviewed, from the initial discovery to current technological breakthroughs and future trends.

Chapter 3 focuses on the research methodology used to conduct this research study. It contains techniques that will be used in the project's design and implementation. It also includes justification for the methods and approaches employed. Chapter 4 focuses on the results obtained from this research. The training