THE PRODUCTION AND PERCEPTION OF ENGLISH SHORT AND LONG VOWELS BY MALAY YOUNG LEARNERS

RAJA ROSEMAWATI BINTI RAJA ABDULLAH

FACULTY OF LANGUAGE AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2021

THE PRODUCTION AND PERCEPTION OF ENGLISH SHORT AND LONG VOWELS BY MALAY YOUNG LEARNERS

RAJA ROSEMAWATI BINTI RAJA ABDULLAH

DISSERTATION SUBMITTED IN [FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGLISH AS A SECOND LANGUAGE

FACULTY OF LANGUAGE AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2021

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Raja Rosemawati binti Raja Abdullah

Matric No: TGB150056/17020783

Name of Degree: Master of English as a Second Language

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"): The Production and Perception of English Short and Long Vowels by Malay Young Learners

Field of Study: Phonetics and Phonology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 01/09/2021

Subscribed and solemnly declared before,

Witness's Signature Date: 2021-09-01

Name:

ABSTRACT

The present study aims to explore how young learners of English as a second language produce English monophthongs and their ability to distinguish the short and long vowels. 10 Year 5 female pupils were selected, and this research focused on three English vowel pairs: /ɪ/ - /iː/, /ʊ/ - /uː/ and /ʌ/ - /ɑː/. The first two research questions looked at the production of these English vowels in terms of formant frequencies and duration. The last research question investigated the ability of the speakers to discriminate English monophthong in terms of accuracy and latency. The data were collected via two types of instruments: the production and perception test. In the production test, the participants were recorded reading six carrier sentences in three repetitions. The findings showed that the /ʊ/ - /uː/ pair was the easiest to produce, with more difference in F1 and F2 values as well as duration difference; while the /ɪ/ - /iː/ pair was the hardest among the three pairs to be produced. In the perception test, participants underwent a listening test with 2AFC. The participants distinguished the short and long vowels, and their accuracy and reaction time were measured. The results indicated that the /ɪ/ - /iː/ pair was the most difficult to discriminate, with more wrong answers recorded and the longer average time was measured. In contrast, the /ʊ/ - /uː/ pair was the easiest to perceive, followed by /ʌ/ - /ɑː/ pair. Comparing the results of both tests, the order of the hardest vowels pair to produce and perceive are as followed; /ɪ/ - /iː/, /ʌ/ - /ɑː/ and /ʊ/ - /uː/. The result agrees with what perception-production link theory has claimed, confirming the relationship between perception and production. The sound that can be perceived easily can also be produced easily. The Speech Learning Model verified that the L2 sounds that are similar to their L1 counterpart are harder to understand and Perceptual Assimilation Model suggested that the L2 sound is assimilated to the most similar sound in L1. These findings help us to understand how Malay young learners learned English vowels, by basing the L2 to L1 (Malay) vowels.

Keywords: production, perception, perception-production link, English monophthongs.

ABSTRAK

Kajian ini bertujuan untuk meneroka bagaimana pelajar muda bahasa Inggeris sebagai bahasa kedua menghasilkan monophthong bahasa Inggeris dan keupayaan mereka untuk membezakan vokal pendek dan panjang. 10 murid perempuan Tahun 5 dipilih, dan kajian ini memberi tumpuan kepada tiga pasangan vokal Inggeris: /ɪ/ - /i:/, /ʊ/ - /u:/ dan /ʌ/ - /ɑ:/. Dua soalan penyelidikan pertama meneliti penghasilan vokal Bahasa Inggeris daripada segi frekuensi dan jangka masa. Soalan penyelidikan terakhir menyiasat keupayaan penutur untuk membezakan monophthong bahasa Inggeris dari segi ketepatan dan latensi. Data dikumpulkan melalui dua jenis kaedah: ujian penghasilan dan persepsi. Dalam ujian penghasilan, para peserta telah dirakamkan membaca enam ayat-pembawa dalam tiga pengulangan. Hasil dapatan menunjukkan bahawa pasangan /ʊ/ - /u:/ adalah yang paling mudah untuk menghasilkan, dengan perbezaan nilai F1 dan F2 serta perbezaan masa yang tinggi; manakala / ɪ / - / i: / pasangan adalah yang paling sukar dihasilkan di antara tiga pasangan vokal dikaji. Dalam ujian persepsi, peserta menjalani ujian mendengar dengan 2AFC. Peserta membezakan vokal pendek dan panjang dan ketepatan masa serta tindak balas mereka diukur. Keputusan menunjukkan bahawa pasangan /ɪ/ - /i:/ adalah yang paling sukar untuk mendiskriminasikan, dengan lebih banyak jawapan salah dicatatkan dan masa purata yang lebih panjang diukur. Sebaliknya, pasangan /ʊ/ - /u:/ adalah yang paling mudah untuk dibezakan, diikuti oleh pasangan /ʌ/ - /ɑ:/. Membandingkan keputusan kedua-dua ujian terbabit, urutan pasangan vokal paling sukar untuk dihasilkan dan dibezakan adalah seperti yang diikuti; /ɪ/ - /i:/, /ʌ/ - /ɑ:/ dan /ʊ/ - /u:/. Keputusan ini bertepatan dengan teori *perception-production link*, mengesahkan hubungan antara persepsi dan produksi. Bunyi yang dapat dibeza dengan mudah juga boleh dihasilkan dengan mudah. *Speech Learning Model* mengesahkan bahawa bunyi L2 yang serupa dengan L1 mereka lebih sukar untuk difahami dan *Perceptual Assimilation*

Model mencadangkan bahawa bunyi L2 diasimilasikan kepada bunyi yang paling serupa dalam L1. Penemuan ini dapat membantu kita memahami bagaimana para pelajar muda Melayu belajar huruf vokal bahasa Inggeris, dengan melandaskan vokal L2 kepada vokal L1 (Melayu).

Kata kunci: production, perception, perception-production link, English monophthongs.

ACKNOWLEDGEMENTS

Alhamdulillah, all praise to Allah for His blessing I am able to complete my study here. I would like to express my profound gratitude to my supervisor, Dr Chiew Poh Shin for her endless support and guidance, throughout my journey in completing my dissertation. The advice, corrections, suggestion and moral supports really get me going. I am also grateful and thankful to be blessed with a supportive family, my mother Normah binti Hamzah, my husband Muhammad Ashraf bin Abdullah, my sisters and others. I will repay and will not forget your assistance, financially and emotionally. This five-year master journey is not easy without encouragement and support from my classmates, fellow CPS supervisees and colleagues. The discussion, suggestion and assistance we had made this dissertation even sweeter. To my cooperative participants from Sekolah Kebangsaan Desa Setapak, teachers and administrators, thank you for allowing me to conduct my research in your school. The data collected is valuable and precious.

May this challenging journey brings me to another level, and may I have the strength to go even further. Thank you, everyone.

TABLE OF CONTENTS

Abstract iv

Abstrak vi

Acknowledgements viii

Table of Contents viii

List of Figures xii

List of Tables xiii

List of Symbols and Abbreviations xvi

List of Appendices xvii

CHAPTER 1: INTRODUCTION 1

1.1 Background of the Study 1
1.2 Statement of the Problem 3
1.3 Research Purpose and Objectives 4
1.4 Research Questions 5
1.5 Significance of the Study 5
1.6 Scope and Limitations of the Study 6
1.7 Organisation of the Study 6
CHAPTER 2: LITERATURE REVIEW 6
2.1 The Vowel System of Malay and English Language 7
2.1.1 Vowel system of Standard English 7
2.1.2 Vowel system of Standard Malay 8
2.2 English Education in Malaysia: Primary School 10
2.3 Production and Perception of English Vowels by English Learners 16
2.3.1 Research on English Vowel Production of English Language 17
2.3.2 Research on English Vowel Perception of English Language 20
2.3.3 Research on Production and Perception of English Language 22
2.4 Theoretical Framework of Vowel Production and Vowel Perception 24
2.4.1 Perceptual Assimilation Model (PAM) 24
2.4.2 Speech Learning Model (SLM) 26
2.4.3 Response Latency 27
2.4.4 Production-Perception Link 28
2.5 Conclusion 29
CHAPTER 3: RESEARCH METHODOLOGY 31
3.1 Participants 31
3.2 Consent 32
3.3 Data Collection 32

3.3.1 Language Profile Test 33
3.3.2 Production Test 33
3.3.3 Perception Test 34
3.4 Procedure 36
3.4.1 Selection of Participants 37
3.4.2 Production Test 37
3.4.3 Perception Test 39
3.5 Data Analysis 41
3.5.1 Language Profile Test 41
3.5.2 Production Test 42
3.5.3 Perception Test 44
3.5.4 Statistical Analysis 44
3.6 Conclusion 45
CHAPTER 4: FINDINGS AND DISCUSSION 46
4.1 Production Test 46
4.1.1 Vowel Quality Contrast 46
4.1.1.1 /ɪ/ and /iː/ Pair 47
4.1.1.2 /ʌ/ and /ɑː/ pair 48
4.1.1.3 /ʊ/ and /uː/ Pair 49
4.1.1.4 Analysis of the Findings 50
4.1.2 Vowel Duration Contrast 53
4.1.2.1 /ɪ/ and /iː/ Pair 54
4.1.2.2 /ʌ/ and /ɑː/ Pair 56
4.1.2.3 /ʊ/ and /uː/ Pair 57
4.1.2.4 Analysis of the Findings 58
4.1.3 Summary 59

4.2 Perception Test 59

4.2.1 Correct Identification 60

4.2.1.1 /ɪ/ and /i:/ Pair 61

4.2.1.2 /ʌ/ and /ɑ:/ Pair 62

4.2.1.3 /ʊ/ and /u:/ Pair 63

4.2.1.4 Analysis of the Findings 63

4.2.2 Reaction Time (Latency) 66

4.2.2.1 /ɪ/ and /i:/ Pair 68

4.2.2.2 /ʌ/ and /ɑ:/ Pair 69

4.2.2.3 /ʊ/ and /u:/ Pair 71

4.2.2.4 Analysis of the Findings 72

4.2.3 Summary 80

4.3 Discussion 81

4.3.1 Production Test 81

4.3.2 Perception Test 86

4.3.3 Relationship between Production and Perception of English Vowels 90

4.4 Conclusion 94

CHAPTER 5: CONCLUSION 94

5.1 Summary of the Findings 95

5.1.1 Research Question 1: What are the qualities of the vowel produced by Malay young learners of English in terms of formant frequencies (F1 and F2)? 95

5.1.2 Research Question 2: To what extent is the difference in duration between /ɪ/ - /i:/, /ʊ/ - /u:/, and /ʌ/ - /ɑ:/ vowels produced by Malay young learners of English? 96

5.1.3 Research Question 3: What is the ability of the Malay young learners in distinguishing English short and long vowels? 97

5.1.4 Discussion on Perception and Production of English Vowels 98

5.2 Implication 99

5.3 Recommendations 100

5.4 Conclusion 101

References 103

Appendix A 115

Appendix B 116

Appendix C 118

Appendix D 121

LIST OF FIGURES

Figure 2.1 English Vowel Cardinal Chart (adapted from Gurnam & Suthagar (2010)) .. 8

Figure 2.2 Malay Vowel Chart (adapted from Clynes & Deterding, 2011)) 9

Figure 2.3 CEFR Target for Each Stage of Education 13

Figure 2.4 Word Recognition Scheme (Adapted from Schiepers, 1980) 28

Figure 3.1 Screenshot of the interface of Psychopy software 40

Figure 3.2 Screenshot of the TextGrid in Praat software for the work 'pick' 43

Figure 4.1 Vowel Distribution for /ɪ/ and /i:/ 47

Figure 4.2 Vowel Distribution for /ʌ/ - /ɑ:/ 48

Figure 4.3 Vowel Distribution for /ʊ/ and /u:/ 49

Figure 4.4 Participants' Vowel Distribution 50

Figure 4.5 Vowel Plotting of the Female Young Learners 52

Figure 4.6 Vowel Duration for /ɪ/ and /i:/ 54

Figure 4.7 Vowel Duration of /ʌ/ and /ɑ:/ 56

Figure 4.8 Vowel Duration of /ʊ/ and /u:/ 57

Figure 4.9 Mean Vowel Duration and Difference in Duration (msec) 58

Figure 4.10 Percentage of Participants' Response 60

Figure 4.11 Percentage of Responses for /ɪ/ and /i:/ Pair 61

Figure 4.12 Percentage of Responses for /ʌ/ and /ɑ:/ pair 62

Figure 4.13 Percentage of Responses for /ʊ/ and /u:/ Pair 63

Figure 4.14 Percentage of Responses for Short Vowel..... 64

Figure 4.15 Percentage of Responses for Long Vowel 65

Figure 4.16 Percentage of Responses for Short and Long Vowels by Participants 66

Figure 4.17 Percentage of Responses for Short and Long Vowel 66
Figure 4.18 Average Reaction Time (s) for /ɪ/ and /i:/ Responses 69

Figure 4.19 Average Reaction Time (s) for /ʌ/ and /ɑ:/ Responses 70

Figure 4.20 Average Reaction Time (s) for /ʊ/ and /u:/ Responses 71

Figure 4.21 Total and Average Reaction Time (s) based on Vowel. 73

Figure 4.22 Total and Average Reaction Time based on Vowel Pairs 75

Figure 4.23 Short Vowels Total Reaction Time (s) by Participants 76

Figure 4.24 Long Vowels Total Reaction Time (s) by Participants 78

Figure 4.25 Total and Average Reaction Times of Correct and Incorrect Responses for Short and Long Vowels (s) 80

Figure 4.26 Vowel Plotting Malay Young Speakers vs Native Speakers (Deterding, 1997) .....
82

Figure 4.27 Relationship of Accuracy and Latency (by vowel) 88

Figure 4.28 Relationship of Accuracy and Latency (by vowel pair) 89

LIST OF TABLES

Table 2.1 English Vowel Pairs Realisation into Single Vowel 9

Table 2.2 KSSR Content Standard and Learning Standard (before CEFR) focusing on pronunciation..... 11

Table 2.3 KSSR Content Standard, Focus and Learning Standard (after CEFR) focusing on pronunciation..... 14

Table 2.4 Perceptual Assimilation Model (PAM) Categories (Komurcu & Yildiz, 2011; Zahariah, 2005) 25

Table 3.1 Word List for Production Test 34

Table 3.2 Word List for Perception Test 34

Table 3.3 The value of duration (msec), F1 and F2 (Bark) of the stimuli 36

Table 3.4 Participants' Language Background 41

Table 4.1 Mean Value for F1 and F2 of Targeted Vowel in Hz and Bark, Euclidean Distance in Bark, and Standard Deviation (in parenthesis) 51

Table 4.2 Duration of the English Monophthong Vowel by Female Malay Young Speakers (in msec) and Standard Deviation (in parenthesis) 53

Table 4.3 /ɪ/ and /i:/ Vowel Duration and Mean Different (msec) 55

Table 4.4 /ʌ/ and /ɑ:/ Vowel Duration and Mean Different (msec) 57

Table 4.5 /ʊ/ and /u:/ Vowel Duration and Mean Different (msec) 58

Table 4.6 Average Reaction Times (in second) of English Monophthong based on Response Accuracy and Standard Deviation (in Parenthesis) 67

Table 4.7 Total Average Reaction Time, Different in Total Average Reaction Time, Average Reaction Time and Different in Average Reaction Time of the Targeted Vowels (s) 74

Table 4.8 Participants' Total Reaction Time (s) and Different in Reaction Time for Short Vowels 77

Table 4.9 Participants' Total Reaction Time (s) and Different in Reaction Time for Long Vowels 79

Table 4.10 Average value of F1 & F2 (Bark), average duration (msec) and rank on production difficulties of the targeted vowels 86

Table 4.11 Reaction Time to Response 88

Table 4.12 Result of Perception Test 90

Table 4.13 Findings of Production Test and Perception Test 91

Table 5.1 Comparison between Vowel Pairs on the Result of Production and Perception Test 99

LIST OF SYMBOLS AND ABBREVIATIONS

2AFC	:	2 Alternative Forced Choice
C	:	Consonant
CG	:	Category Goodness
CVC	:	Consonant-vowel-consonant
D	:	Distractor
ED	:	Euclidean Distance
EFL	:	English as a Foreign Language
ESL	:	English as a Second Language
F1	:	First formant
F2	:	Second Formant
Hz	:	Hertz
L1		first language
L2	:	second language
M	:	Mean
msec	:	millisecond
PAM	:	Perceptual Assimilation Model
RP	:	Receive Pronunciation
s	:	second
SC	:	Single Category
SD	:	Standard Deviation
SLM	:	Speech Learning Model
SSB	:	Standard Southern British English
TC	:	Two Category
TW	:	targeted word

LIST OF APPENDICES

Appendix A: Surat Kelulusan Bersyarat untuk Menjalankan Kajian 116

Appendix B: Letter of Consent from Parent/Guardian 117

Appendix C: The Language Profile Questionnaire 119

Appendix D: List of Target Words and Images for Perception Test 122

CHAPTER 1: INTRODUCTION

This research is conducted to investigate the ability of second language speakers in producing and perceiving English monophthong. This chapter presents the background of the research, problem statement, research purpose, objectives and questions, the significance of the research as well as the scope and limitation faced. The organization of this research is also explained in this chapter.

1.1 Background of the Study

Communication can happen both verbal and written. Verbal communication is used more widely in our daily life, either in an informal or formal situation. In order for communication to be effective, the transmitter's message must be understood by the receiver and one of the important elements of effective communication is language use. Communicate in one's own native language may not be a problem, but the issue arises when one has to communicate in other languages, for example, English. English came to Malaysia through the colonization of Malaya before the independence. The language is now the official second language of this country. Thus, it is safe to assume every Malaysian is able to communicate in English, whether verbally or in writing. Unfortunately, many people do not understand that language learning is not about the amount of linguistic knowledge acquires but also a skill to be utilized. The language skills comprise of listening, speaking, reading and writing. However, Malaysians, just like any other Asian community, focuses more on academic performance rather than the actual skill acquire. In acquiring the language, grammar is priorities as it is considered the hardest part of the language. Apart from grammar, pronunciation is considered hard to acquire but yet the teaching of proper or correct pronunciation is neglected. Mispronunciation could lead to misunderstanding and embarrassment (Shak, Lee, &

Stephen, 2016). Prodanovska-Poposka (2017) in her study revealed that ESL (English as Second Language) students were having trouble communicating with English native speakers as some words were incorrectly pronounced and this led to ambiguity. They become uncomfortable speaking in English which causes many to avoid communicating in English as they felt ashamed. The ideology of speaking English with more accurate pronunciation shows that one is more professional and has higher social status contribute to this issue.

The Second Shift of the Malaysian Education Blueprint 2013-2025 emphasizes on every child language proficiency, both Malay and English, as well as encouragement for foreign languages ("Malaysia Education Blueprint 2013 - 2025," 2013). English is the second language in Malaysia, frequently used in formal communication and the international language of communication. Most Malaysian learnt English for 11 years; 6 years in primary school and 5 years in secondary school. English is learnt from various sources, but the language is properly learnt in school. Historically, Malaysian students are excellent in the content-producing subject, but opposite in terms of practicality. Their literacy is highly conceptualized through school experience, only practice the language during English class (Elya & Parilah, 2015). Due to stress on passing the examination, the teaching of speaking, specifically on pronunciation normally is neglected since this part of English is considered unimportant or difficult to teach (Rajadurai, 2006). This resulting in students who are good in English, but not able to use and practice the language.

The effect of negligent teaching pronunciation shows during the tertiary level education and working life. To enrol on degree programmes, Malaysian students are required to seat for English proficiency tests such as MUET (Malaysian University English Test) or IELTS (International English Language Test System). This is because many colleges and universities practice English as the medium of instruction, even for non-English major programmes. Both of these tests evaluate the students on the four

language skills, and speaking part of these two tests do test on the pronunciation, as stated in their rules and regulations (IELTS, n.d.; Malaysian Examinations Council, 2016). How can one be evaluated on something they never learnt? This is unfair but no one really put up the questions. Once they passed this test, the students are able to pursue their studies in any college or university. Now, not only the written assignments are to be prepared in English, oral assignments such as presentations are also to be done in English. Even though the correct pronunciation is not part of the marking scheme but being able to present properly and smoothly, correct pronunciation influences the students' confidence and self-esteem (Soozandehfar, 2012).

After finishing tertiary level education, these graduates are now looking for jobs. Unfortunately, many fail to land a good job, as employers find their communication skills, specifically in English is not up to par (Idham, Yusof, Syazana, Effa, & Talib, 2014). Bad pronunciation can affect communication. Amutan, Ramalingam, Hee, & Maruthai (2017) highlighted that some of the features identified by Human Resource during the interview were poor utterance, poorly constructed sentences and poor communication skills. Hence, the problem occurs when they are applying for a job in the future, as the demand for employees to have a high level of English proficiency is increasing ("Employers : Fresh graduates have unrealistic expectations," 2015), and news such as Faiz (2018), "Malaysian graduates face job interview jitters, no thanks to weak English skills" (2015), "Why many graduates can't express themselves in English" (2017) and Sani (2019) were also reporting similar cases.

1.2 Statement of the Problem

Speaker intelligibility is important to ensure the message is properly delivered. Intelligibility is defined as to what extent the speech can be understood (Lim & Nazri, 2013). Thus, the inability to produce or perceive correct sounds will greatly affect communication, but no specific method to teach pronunciation caused problems among

English teachers in Malaysia. Ahmadi & Abbas (2011) claimed that pronunciation is the least favourite topic to be taught in the classroom. As one of the steps to improve education quality, the Ministry has introduced KSSR to replace KBSR. The Kurikulum Standard Sekolah Rendah (KSSR) was introduced in 2011, aimed to holistically expand pupils' potential in order to produce a balance, harmony, creative, innovative and as preparation to face 21st-century challenges. This curriculum is designed to help the pupils to be more proficient in this English language, with the goals of assisting them to be able to use the language not only for daily life but for future work purposes.

Recently, the ESL and EFL (English as a Foreign Language) community is opened up to variances of the English language, especially towards to local English variant. In contrast to this movement, the Malaysian education system seems to be relying on British English or Receive Pronunciation (RP). RP is the recommended English variant for foreigners who study the English language (Roach, 2007), though it is not the official English accent. Due to Malaysian English is based strongly on the RP; hence the way of pronouncing English words is similar. However, does it really necessary for Malaysians to speak English as to how it is spoken by a native speaker? Jayapalan & Pillai (2011) found that this expectation is unrealistic, especially since the RP is not internationally or world-widely use. Moreover, many ESL and EFL teachers were never studied under a native speaker (Rajadurai, 2006) and the opportunity to encounter and communicate with native speakers are also low. Hence, this research is looking at how Malay young speakers speak and interpret the English sounds, whether following the English standard or not.

1.3 Research Purpose and Objectives

The purpose of this research is to explore how young learners of English as a second language produce English short and long vowel sounds and their ability to distinguish the sounds. Thus, the objectives of this research are:

a) to examine the production of English short and long vowel sounds by Malay young learners.

b) investigate the perception of English short and long vowels by Malay young learners.

1.4 Research Questions

The research questions for this research are as follow:

1. What are the qualities of the vowel produced by Malay young learners of English in terms of formant frequencies (F1 and F2)?

2. To what extent is the difference in duration between /ɪ/ - /i:/, /ʊ/ - /u:/, and /ʌ/ - /ɑ:/ vowels produced by Malay young learners of English?

3. How do the Malay young learners of English perceive the English short and long vowels?

The first and second questions look at the sound production by the participants. The sounds are analyzed in terms of duration (Q1) and vowel quality (Q2). As for the third question, participants will undergo a test (like a listening test) and must distinguish the given sounds.

1.5 Significance of the Study

This research is hoped that it will contribute to the body of knowledge of production and perception of English vowels produced by Malay speakers, especially the young speakers. This is because the number of researches on this topic focusing on young speakers is still lacking. Other than that, this field is lacking researches that investigate both production and perception of the target language. The findings of this research could help to supply better insight into Malaysian students' pronunciation which could lead to aid teachers, especially those who are teaching primary school in understanding their students' capability in distinguishing sounds, in this case, English monophthongs. Thus, it could be beneficial in improving English proficiency among Malaysians.

1.6 Scope and Limitations of the Study

The present research is looking at the ability to distinguish English short and long vowels by Malaysian young speakers. There are twelve English monophthongs, but this research is focusing on the three vowel pairs only, which are /ɪ/ - /i:/, /ʊ/ - /u:/, and /ʌ/ - /ɑ:/. Thus, the result cannot be generalised to the other monophthong, or other vowels. Moreover, the participants of this research are the Malay young learners; specifically, females, selected from Year 5 pupils of Sekolah Kebangsaan Desa Setapak. This hence it cannot be generalized to the other categories. However, this research can be adapted to different categories such as age, gender and education level in order to understand their ability in producing and perceiving English sounds.

1.7 Organisation of the Study

This research contains five chapters. The first chapter explains the problem and purpose of the research. The second chapter discusses the framework of the research and the related literature for better understanding. Next, chapter three explains the methodology applied in order to obtain the data, and also how the data is to be analysed. The findings of the research are presented in chapter four, and some discussions are also conducted here. Finally, chapter five will answer the research questions and summarise this research.

CHAPTER 2: LITERATURE REVIEW

In order to understand this research, this chapter will discuss some literature involved in the perception and production of English vowels. The chapter compares and contrasts the Malay vowel system and English vowel system, providing some insight on how the participants may respond to the stimuli prepared. The teaching of the English language in Malaysia is also discussed to provide some expectations. Several theories on the perception and production of a second language are also discussed. Past research is also

studied to provide better background knowledge and discussion. Lastly, the gap in this field is discovered and discussed, hence the relevance of this research.

2.1 The Vowel System of Malay and English Language

Most language learners dreamt to sound like native speakers, however, achieving this is quite difficult. The least they could do is to be able to pronounce the words as clearly as possible. In terms of reading, writing or even grammar, most learners can understand and practice them correctly. However, the case is somehow different in speaking. There are various factors contributing to the inability to speak native-like, such as first language interference, generalization, fossilization, and others. To speak a language properly, one must understand the basis of sound and the sound production of the spoken language, both native and foreign languages.

2.1.1 Vowel system of Standard English

Malay and English may seem like they are having similar vowel sets but unfortunately, they are not. Various researches were conducted to identify the appropriate F1 and F2 of the English vowels and this research chose Deterding (1997) as the reference. The participants of this research speak in Standard Southern British English (SSB), also known as the Received Pronunciation (RP). Malaysian English is based on the RP, hence the way of pronouncing English words is similar. English has 11 monopthongs; six short vowels (/ɪ/, /ʊ/, /ʌ/, /ɒ/, /ə/, /e/) and five long vowels (/iː/, /uː/, /ɑː/, /ɔː/, /ɜː/), and eight

English diphthongs; /aɪ/, /eɪ/, /əʊ/, /aʊ/, /eə/, /ɪə/ and /ɔɪ/ (Gurnam & Suthagar, 2010; Roach, 2007). Figure 2.1 illustrate the standard English vowel chart and the positions of each vowel. Starting at the top-right area is the location of /iː/, follows by /ɪ/, positioned lower than /iː/ and more towards the centre. /uː/ is aligned to /iː/ as it is also on the top but a way to the back of the chart. The short version, /ʊ/ is aligned with /ɪ/, also closer to the centre but placed on the left side of the chart. Lastly, ʌ/ is at the centre of the chart and

the lower area. It is positioned higher than its counterpart /ɑ:/, which positioned at the lowest of the chart, toward its back, slightly more back than /u:/. In general, the long monophthongs are placed outer than their short monophthongs.

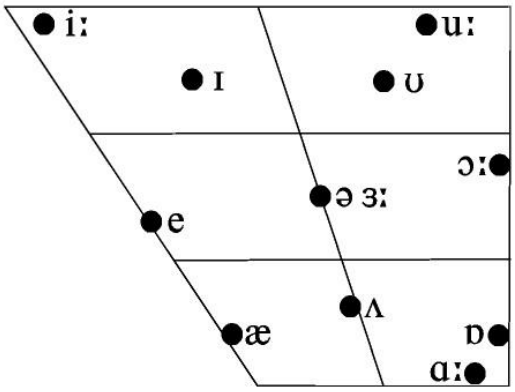

Figure 2.1 English Vowel Cardinal Chart (adapted from Gurnam & Suthagar (2010))

2.1.2 Vowel system of Standard Malay

Both Malay and English languages have monophthong and diphthong vowels, yet the idea of vowel duration does not exist in the Malay language. Malay vowel system consists of six monophthong vowels; /i/, /u/, /a/, /ə/, /e/ and /o/, and three diphthong vowels; /ai/, /au/, and /oi/ (Nik, Farid, Hashim, & Hamid, 2008). Figure 2.2 shows the Malay vowel chart and the position of each vowel.

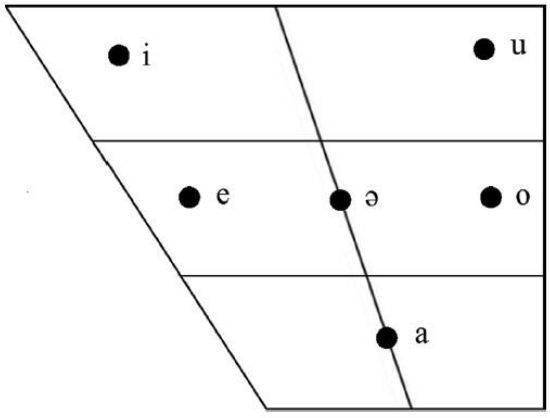

Figure 2.2 Malay Vowel Chart (adapted from Clynes & Deterding, 2011))

The /i/ in Malay is located at the front and high, similar to /iː/ and /ɪ/ in English. This applies to /u/ also as it is situated at back and high in Malay, just like /ʊ/ and /uː/ in English. However, /a/ in Malay is positioned at the front and low, and this position is more similar to English /ʌ/ than /ɑː/, as it is located at mid and low while /ɑː/ is at back and low. In addition, the /ə/ in Malay is similar to English /ə/ rather than the long version /ɜː/. Anyhow, the position of /o/ in Malay is more similar to /ɔː/ than the short version /ɒ/. the. Lastly, Malay /e/ is slightly inward than the English counterpart. This proves that it is not reasonable to assume the Malay vowels are short vowels and similar to English short vowels, but study Zuraidah (1997) (cited in Pillai, Don, Knowles, & Tang (2010)) found that the English vowel pairs were realised as a single vowel as illustrates in Table 2.1.

Table 2.1 English Vowel Pairs Realisation into Single Vowel

English Vowel	Realised as Single Vowel
[ɪ] and [iː]	[i]
[ʊ] and [uː]	[u]
[e] and [æ]	[e]
[ɒ] and [ɔ]	[o]
[ʌ] and [ɑː]	[a]
[ə] and [ɜ]	[ə]

Despite the differences, Malaysian Malays learnt English as their second language for 11 years, during compulsory primary and secondary school. Most Malay can perform excellently in reading and writing but not for listening and speaking. The ability to both produces and perceives the targeted sound properly is important in communication. As for English, mispronunciation of a sound could change the meaning of the word, hence changing the entire direction of the message. When the structure of a second language (L2) does not exist or is different than in the first language (L1), the L2 learners will have difficulties in understanding the aspect (Norsimah, Kesumawati, Norzakiah, & Norhashimah, 2007). The absence of vowel quality contrast and length discrimination did hinder them to communicate effectively. This typifies their low proficiency in the English language, which could lead to a secure proper job in future (Nasrin, 2018).

2.2 English Education in Malaysia: Primary School

Malaysian education has undergone changes, with the implementation of KSSR for a primary school in 2011. Our examination system is well-known to be result-oriented, where students and parents put lots of emphasis on the marks and grades. Due to this, spoken language faced discrimination as less emphasis was given. The teaching of pronunciation is always neglected (Jayapalan & Pillai, 2011). This action is considered typical among English teachers for whom English is not their native language. Thus, a new curriculum was introduced and the learning standards do focus on pronunciation throughout the schooling years. The teaching of phonics was introduced to the pupils as early as Year 1 ("Malaysia Education Blueprint 2013 - 2025," 2013) and Table 2.2 shows the content standard and learning standard of KSSR, only highlighting those associated with teaching pronunciation. The teaching of pronunciation is covered in three of the teaching modules, listening and speaking, reading, and language arts. This syllabus allows the students to experience language learning as a whole. Their perception of

English sounds is trained early, and the production of the sounds continues throughout their study years. The table shows that the perception of language is introduced first in

Year 1 and Year 2, before working on the production of the sounds, which starts at Year 3. After two years of implementation, this curriculum is revised with the introduction to CEFR.

Table 2.2 KSSR Content Standard and Learning Standard (before CEFR) focusing on pronunciation

Content Standard	Learning Standard					
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Listening and Speaking Skills						

1.1 By the end of the 6-year primary schooling, pupils will be able to pronounce words and speak confidently with the correct stress, rhythm and intonation.	1.1.1 Able to listen and respond to stimulus given with guidance: a) environmental sounds b) instrumental sounds c) body percussion d) rhythm and rhyme e) alliteration f) voice sounds g) oral blending and segmenting	1.1.1 Able to listen and respond to stimulus given with guidance: a) environmental sounds b) instrumental sounds c) body percussion d) rhythm and rhyme e) alliteration f) voice sounds g) oral blending and segmenting	1.1.1 Able to speak with correct word stress.	1.1.1 Able to speak with correct word stress.	1.1.1 Able to speak with correct pronunciation, stress, rhythm and intonation.	1.1.1 Able to speak with correct pronunciation, stress and intonation
			1.1.3 Able to listen to,	1.1.3 Able to listen to	1.1.3 Able to listen to	1.1.2 Able to listen to
			say aloud and recite rhymes, tongue twisters and sing songs paying attention to pronunciation, rhythm and intonation.	and recite poems, tongue twisters and sing songs, paying attention to pronunciation.	and respond to a given stimulus by using appropriate words, phrases and expressions with the correct stress, rhythm and intonation	and respond confidently to a given stimulus by using appropriate words, phrases and expressions with the correct stress and intonation

Reading Skill						
2.1 By the end of the 6-year primary schooling, pupils will be able to apply knowledge of sounds of letters to recognise words in linear and non-linear texts.	2.1.2 Able to recognise and articulate initial, medial and the final sounds in singlesyllable words within given context:	2.1.1 Able to recognise and articulate initial, medial and the final sounds in singlesyllable words within given context:				
	2.1.3 Able to blend two to four phonemes into recognizable words and read them aloud.	2.1.2 Able to blend phonemes into recognizable words and read them aloud.				
Language Art						
4.1 By the end of the 6-year primary schooling, pupils will be able to enjoy and appreciate rhymes,	4.1.2 Able to recite nursery rhymes, jazz chants and sing action songs with correct pronunciation and rhythm.	4.1.2 Able to sing action songs and recite jazz chants with correct pronunciation, rhythm and intonation.	4.1.2 Able to sing action songs recite jazz chants and poems with correct pronunciation, rhythm and intonation.	4.1.2 Able to sing songs and recite jazz chants and poems with the correct stress, pronunciation, rhythm	4.1.2 Able to listen to, sing songs, recite jazz chants and poems with the correct stress, pronunciation, rhythm and intonation	4.1.2 Able to sing songs and recite jazz chants and poems with the correct stress, pronunciation, rhythm and intonation
poems and songs, through performance.				and intonation.		

In 2013, the Common European Framework of Reference for Languages or also known as CEFR was introduced to the syllabus, in alignment with the establishment of English

Language Standards and Quality Councils (ELSQC). However, the curriculum was only fully implemented in 2017. Figure 2.3 shows the target for each stage of education according to CEFR. In primary school, pupils are expected to achieve at least A2, in which they can communicate in a simple and routine task. The target allows the teacher to plan their lesson properly in order to achieve it, as well as to challenge their students. After all, the students are being prepared to use the language in their daily life.

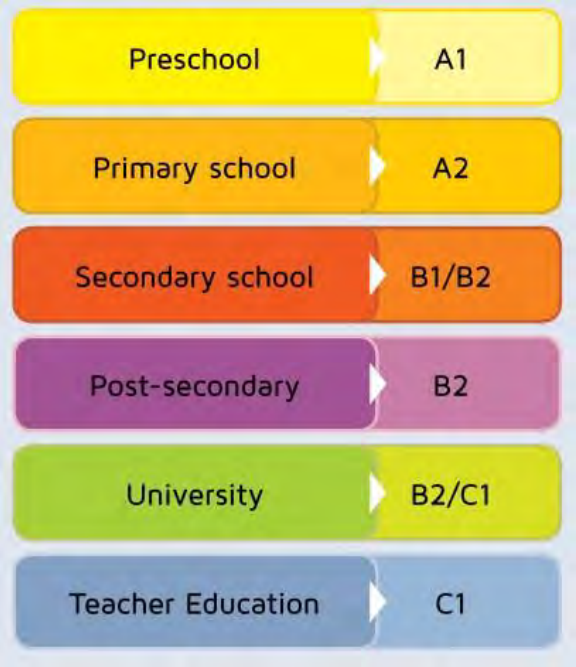

Figure 2.3 CEFR Target for Each Stage of Education

It would be impractical to teach phonetics and phonology to young learners, thus, the teaching of phonemic and phonological awareness was introduced in 2011. According to the English Language Standards and Quality Council (2015), the lesson starts by enhancing language perception. Beginner learners must understand the concept of phonemes (or phonics rules) and gradually, this will lead to the production of the words; both in written and spoken. Teaching correct pronunciation and intelligibility in communication are introduced as early as Year 1. Table 2.3 shows the syllabus for primary school; the content standard, focus and learning standard that emphasises production and perception of the language. The additional 'focus' is to help teachers in

planning their teaching better. Also, listening and speaking are now in separate modules. Apart from the writing module, the other language modules do highlight and teach the students phonemic and phonological awareness, including Language Arts.

Table 2.3 KSSR Content Standard, Focus and Learning Standard (after CEFR) focusing on pronunciation.

	Learning Standard					
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Listening Skills						
Content Standard	1.1 Recognise and reproduce target language sounds					
Focus	Recognise and reproduce target language phonemes intelligibly					
	1.1.1 Recognise and reproduce with support a limited range of highfrequency target language phonemes	1.1.1 Recognise and reproduce with support a range of highfrequency target language phonemes	1.1.1 Recognise and reproduce with support a range of target language phonemes	1.1.1 Recognise and reproduce with support a wide range of ta rget language phonemes	1.1.1 Recognise and reproduce with little or no support a wide range of target language phonemes	1.1.1 Recognise and reproduce independentl y a wide range of target language phonemes
Speaking Skills						
Content Standard	2.1 Communicate simple information intelligibly					
Focus	Communicate simple information about themselves clearly					
	2.1.1 Give very basic personal informatio n using fixed phrases	2.1.1 Give simple personal informatio n using basic statements	2.1.1 Ask about and express basic opinions	2.1.1 Explain and give reasons for basic opinions	2.1.1 Give detailed informatio n about themselves	2.1.1 Give detailed information about themselves and others
Focus	Find out simple information from others					

	2.1.2 Find out about very basic personal information using fixed phrases	2.1.2 Find out about personal information by asking basic questions	2.1.2 Find out about and describe basic everyday routines	2.1.2 Find out about and describe experiences in the past	2.1.2 Find out about and describe experiences up to now	2.1.2 Ask about and express rules and obligations
Focus	Communicate simple information clearly					
	2.1.3 Express basic likes and dislikes	2.1.3 Give a short sequence of basic instructions	2.1.3 Give a short sequence of basic directions	2.1.3 Give a longer sequence of basic instructions or directions	2.1.3 Ask for, give and respond to simple advice	2.1.3 Explain and give reasons for simple advice
	2.1.4 Greet, say goodbye, and express thanks using suitable fixed phrases	2.1.4 Ask about and express ability	2.1.4 Ask about, make and respond to simple predictions	2.1.4 Give reasons for simple predictions	2.1.4 Ask about and describe future plans	2.1.4 Ask about and describe future plans or events
Focus	Describe people and things clearly					
	2.1.5 Name or describe objects using suitable words from word sets	2.1.5 Describe objects using suitable words and phrases	2.1.5 Describe people and objects using suitable words and phrases	2.1.5 Describe people, and objects using suitable statements	2.1.5 Describe people, places and objects using suitable statements	2.1.5 Ask about and describe personality
Reading Skills						
Content Standard	3.1 Recognise words in linear and non-linear texts by using knowledge of sounds of letters					
Focus	Distinguish and articulate beginning, medial and final sound words					
	3.1.2 Recognise and sound out with support beginning, medial and final sounds in a word	3.1.2 Recognise and sound out with some support beginning, medial and final sounds in a word				
Focus	Blend phonemes to recognise words					

	3.1.3 Blend phonemes	3.1.3 Blend phonemes (CVC,		
	(CVC, CCVC)	CCVC, CVCV, CCV		
Focus	Segment words into phonemes to spell			
	3.1.4 Segment phonemes (CVC, CCVC)	3.1.4 Segment phonemes (CVC, CCVC, CVCV, CCV		
Language Art				
Content Standard	5.1 Enjoy and appreciate rhymes, poems and songs			
Focus	Say the words in simple texts, and sing simple songs with intelligible pronunciation, rhythm and intonation			
	5.1.2 i) simple chants and raps ii) simple rhymes iii) simple action songs	5.1.2 In addition to Year 1 text types: simple songs	5.1.2 In addition to Year 2 text types: simple poems	

Comparing the new syllabus with the previous version, more details are provided for each of the learning standards. As now the listening and speaking modules are separated, the perception and production teachings are now clearly defined and planned. The teaching now emphasises the phonemes, including blending the phonemes to make words and segmenting the word to identify phonemes. The speaking module provides various situations and concepts for the students to use the language and get comfortable practising them.

2.3 Production and Perception of English Vowels by English Learners

The study of language production is intriguing, leading to much researches were conducted. Other than production, research on language perception is also gaining more

interest. The connection between language perception and language production is interesting and this motivates various researchers to study them.

2.3.1 Research on English Vowel Production of English Language

Research by Pillai & Delavari (2012) concluded that Persian speakers tend to conflate the vowel pair, even if the vowel existed in the Persian language. This research is interesting as the participants, 13 Iranian English speakers, were tested in two situations, Word List Context (WLC) where they read the given carrier text in naturalistic context and Informal Speaking Context (ISC) where they talked informally about their life. The findings showed that the participants conflated /ʌ/ and /ɒ/ (except for male WLC), due to the fact that the Persian language does not have them, instead, the language has /o/. Furthermore, the quality contrast between the vowel pairs was also low, with the /i:/ and /ɪ/ were closely produced. The quality of the vowel was differed due to the movement of the lips rounding, which was closer. As for the vowel duration, the participants did produce long vowels longer than the short vowel counterpart, although some pairs have no clear differences. On contrary to the vowel duration finding, another research conducted with Javanese students produced a different result (Perwitasari, Klamer, & Schiller, 2015). The participants were Javanese young adults and American English native speakers. The findings indicated that Javanese speakers produced a shorter duration of the targeted vowels, both short and long vowels. This is because the duration cue does not exist in the Javanese language, causing the speakers to be unaware of this situation. Further study was conducted on the Javanese speakers in terms of English vowel quality contrast, and the findings were compared to Sudanese and native speakers (Perwitasari, Klamer, & Schiller, 2016). All participants were young adults at an average age of 22 years old. They had undergone a minimum of 9 years of learning English formally. Similar to any other L2 speakers, the Javanese and Sudanese English vowel spaces were smaller than the native speakers. Other than that, the vowels were mostly produced more

backward than the native speakers did. Not only they encountered difficulties in producing both new vowels (not exist in their L1), but similar findings were found on similar vowels (exist in L1).

Singaporean speakers were able to differentiate the vowels in terms of duration but facing difficulties in terms of quality (Tan & Low, 2010). This research is interesting as the result was not being compared to native speakers, but other L2 speakers, specifically to Malaysian speakers. Both Singapore and Malaysian English were categorised in the same category as both were colonised by the British for decades. However, Schneider's Dynamic Model of Postcolonial Englishes placed Malaysian English in Phase 3 while Singapore English in Phase 4. Singapore policy of making English their first language enhanced the production of an almost native-like speaker, unlike Malaysia, English is formally introduced in primary school. Surprisingly, the research showed that the vowels were conflated in Singapore English, similar to Malaysian English speakers. English is the official second language in Malaysia. Malaysian English is supposedly based on British English or RP, as we are heavily influenced by the British education system. However, according to (Pillai, 2017), the term Malaysian English is unique as the pronunciation is not exactly similar to Standard British English or the RP. The speakers' L1, education, and socio-economic can contribute to the uniqueness of the sounds produced. The study on Malaysian pronunciation is interesting, thus encouraging various researchers to conduct their research.

Salwani (2005) showed that Malay speakers met difficulties to differentiate between long and short vowels in terms of duration. Surprisingly, the participants of this research were speakers who were majoring in English studies. The research focused on two sets of vowels: /i:/ – /ɪ/ and /u:/ – /ʊ/. Several of the participants produced a longer duration for /ʊ/ than the long version, /u:/. Although there was a significant difference in duration, the

qualities of both short and long vowels were almost similar. Another research that investigated participants who were majoring in the English language was conducted by (Pillai et al., 2010). Indifferently, rather than focusing on one race, this research included other races such as Malay, Chinese, Indian and Eurasian, 47 participants altogether. This research looked at all 11 English monophthongs, embedded in these words; bid, bead, beg, bag, bug, bard, pod, board, put, boot and bird. In the discussion, the researchers mention that "the contrast between vowels in Malaysian English does not match the classical notion of phonemic contrast." The results showed that some pairs were lack of contrast (/i:/ – /ɪ/, /ʌ/ - /ɑ:/ and /e/ - /æ/) while the others presented some realisation (/u:/ – /ʊ/ and /ɒ/ - /ɔ:/). Similar to Salwani's finding, the were no clear quality contrast between the vowel pairs. As for the duration, except for the /ɒ/ and /ɔ:/, the other vowel pairs have a clear significant different. These two types of research showed that event speakers who were majoring in the English language still having difficulties clearly produce each vowel in terms of vowel quality contrast and duration contrast.

In 2007, Pillai conducted an investigation of whether Malaysian English pronunciation is more similar to British English or American English. Generally, students were taught English in reference to RP, but exposure to American English is now wider and regularly due to borderless internet access. She found that the vowel quality of Malaysian English is different compared to American or British English. The participants of this research were taken from three major races in Malaysia, Malay, Chinese and Indian, thus eliminate the ethnicity factor. Malaysian English vowels were more compact in space, a contrast to RP. Although there was a notion that Malaysian English is somehow more similar to American English than British English, Pillai has proven otherwise. Malaysian English is unique on its own.

2.3.2 Research on English Vowel Perception of English Language

Foreign speakers tend to rely on the duration discrimination when identifying an English vowel minimal pair, although they may not produce much difference during the production (Casillas, 2015; Kondaurova & Francis, 2008; Perwitasari et al., 2015). A lot of studies found that there is positive discrimination on the duration of the sound, as longer sounds were produced longer and vice versa. In Kondaurova & Francis (2008), they conducted two different experiments to check how duration cue is used to assist vowel perception English /i:/ and /ɪ/. They found out that L2 speakers utilised duration cues rather than spectrum cues. The finding was compared to native American English speakers, in which this group of participants utilised both spectrum and duration cues. The second experiment of the same research looked at the reliance on duration cue to native phonological contrast, with an unexpected result cited that "Russian, Spanish, and English speakers employ vowel duration to the same degree when identifying stress contrasts in their native languages despite the very different roles that duration plays in the stress systems of the three languages." This indicates that the existent of duration cues in the L1 structure assists them in perceiving L2 sounds. Different research conducted by (Escudero & Boersma, 2004) looked at the speech perception of Spanish speakers and compared it to two different dialects in the English language, Scottish English and Southern English (RP). Again, this research also looked at English /i:/ and /ɪ/ vowels.

"*Scottish English and RP is difference in the production of the /i:/-/I/ contrast for Scottish versus Southern English speakers with respect to the relative use of the acoustic dimensions that signal the contrast. Regarding spectral vowel height, as expressed by the first formant frequency (F1), there is a large height difference between Scottish and Southern English /I/."*

The participants perceived the L2 by reusing their L1 category, as suggested in the

Full Transfer hypothesis (Schwartz & Sprouse 1996 as cited in Escudero & Boersma (2004)). The Scottish /i:/ and /ɪ/ were perceived as /e/ and /i/ in Spanish while the SE /i:/ and /ɪ/ were perceived as Spanish /i/. This hypothesis was confirmed (in this research) and supported by Best's PAM and Fledge's SLM, in which the L2 learners use their L1 knowledge and structure in order to understand the L2.

English as a Foreign Language (EFL) learners may have difficulties in distinguishing the English vowel. Ammar, Ilyana, Che, & Yap (2016) conducted research on how the participants' language proficiency affects their perception of English vowels. The participants of the research were Iraqi speakers who were studying in Malaysia and this research studied all 12 English monophthongs (including /æ/). The result indicated the proficiency did assist them to identify the vowel, however, they still encountered problems in doing so. The long vowels did not pose many challenges as they are also present in Iraqi Arabic, however, the participants were making more errors on the lax vowels. For example, many errors were produced in perceiving /ɪ/ to /ɛ/. EFL learners also have the tendency to assimilate the L2 sound(s) into their L1 sound(s). According to Yang & Fox (2014), monolingual Chinese participants perceived English vowels by assimilating them into their native language vowels. This research studied 13 English vowels; 10 monophthongs and 3 diphthongs, analysing the perceptual structure using perceptual dimensions. The perceptual space of monolingual Chinese listeners is different from the other two listener groups, as they clustered the vowels into three main groups; rounded vowels, spread vowels and slightly spread vowels. Thus, this shrank the perceptual space of English vowels. In the same research, bilingual Chinese speakers did perform better in perceiving the English vowel, in comparison to the monolingual speakers, confirming that exposure and experience can improve one's perception.

Since English is Malaysian's second language, communicating in this language is the norm. The production of the language is important, but perceiving the language is also important. According to Thai, Eng, & Aziz (2010), Malay bilingual speakers were having difficulties in distinguishing between English short and long vowels. The study looked at five English monophthongs, embedded in beat, bit, bait, bet and bat, 52 Malay-English bilingual undergraduates participated in this study. The results showed that the participants assimilated sounds that were not familiar into sound which was more familiar to them (assimilating non-Malay sounds into the closes Malay sounds). For example, they assimilated the English sound /ɛ/, which was not present in their native language to /e/, which does present in Malay. Although this research only focused on the front vowel, this gives a better insight into understanding Malay speakers' perceptions.

2.3.3 Research on Production and Perception of English Language

Research on the /ɪ/ - /i:/ contrast by Spanish speakers indicated that the early learners were able to produce the sound with good contrast; in terms of both duration and quality, but the late learners were not as good as the early learners (Casillas, 2015). The participants of this research were 30 adult females, categorised into early learners of English (EL), late learners of English (LL) and native monolingual of English (NS). They were to recite carrier phrases containing all English vowels but only /ɪ/ and /i:/ were analysed. The EL produced /ɪ/ and /i:/ contrast almost native-like, however, the LL managed to produce /ɪ/ shorter than /i:/, but no distinct contrast on the acoustic analysis. For the perception experiment, the participants were to identify the sheep-ship token. The LL relied solely on the duration cue while EL and NS relied more on the spectrum cue. Early exposure to the target language helps them greatly in their communication. The effects of sequential language are present, but earlier exposure manages to improve language acquisition. However, as for the participants of this research, they were living in an English-speaking society, which allows them to practice the acquired language at

ease. Another research on Spanish speakers by Casillas & Simonet (2016) claimed that these speakers; whether they were exposed to the English language early in their life or later, were still unable to produce the vowel /æ/ and /ɑ/ like the native speakers do. The participants for this research were divided between Spanish native who can be considered bilingual, as they spoke both Spanish and English, and received English language exposure late in their adulthood (LL), and Spanish native who first language was English rather than Spanish (EL) as they were exposed to English very early. English monolingual speakers who were recently being exposed to Spanish were also included in this research. In the production experiment, participants were to recite carrier phrases consisting of various English monophthongs, but only words consist of /æ/ and /ɑ/ were analysed. As for the perception experiments, there were two tests conducted, identification task and discrimination task. Although the EL group claimed to be English-dominant, the findings showed that they behaved similarly to LL in both production and perception task. The acoustic distance was smaller and the acoustic distribution showed some overlapping. However, the EL did better in the discrimination task, producing almost similar result to NE but with more false alarms. Thus, sequential bilingualism can affect language acquisition, even if the foreign speakers are more dominant towards the L2. Unfortunately, these two researches did not make any connection between the ability to perceive and ability to produce second language.

However, Shahidi, Aman, & Kechot (2012) research produced rather an interesting finding. Rather than vowel, the study looked at English stop consonant; /p/, /b/, /t/, /d/, /k/ and /g/. In the production experiment, the participants recited words of the target consonants both in Malay and English settings. As for the perception experiment, the participants underwent an identification test. Their Malay speaking participants somehow did not manage to perceive the targeted sound, although they did produce the same sound properly earlier. Thus, the ability to accurately produce a sound does not

assist the speaker to be able to perceive the sounds accurately too. Very little researches were conducted on the production and perception of the English language by Malay speakers, and even fewer on the English vowels. Most research will either focus on production or perception of the language.

2.4 Theoretical Framework of Vowel Production and Vowel Perception

Isbell (2016) defined perception as a process of aural input, involving the recognition of sounds based on phonological category. On the other hand, production is the emitting of an auditory signal by using the oral-articulatory system, which brings a message that is linguistically encoded. A few theories and models have been proposed and developed, in order to understand the production and perception of language by learners. This research is using the Perceptual Assimilation Model and Speech Learning Model as the main theoretical framework. In addition, the production-perception link is also discussed to make the connection and relevance of the research.

2.4.1 Perceptual Assimilation Model (PAM)

Perceptual Assimilation Model (PAM) introduced by Best (1993, 1994, 1995) (as cited in Best & Tyler, 2007) claimed that the learning of the second language (L2) sounds system is assisted by the first language (L1) sound system (Best, Goldstein, Tyler, & Nam, 2009; Best & Tyler, 2007; Best, Tyler, Goldstein, & Nam, 2016; Tyler, Best, Faber, & Levitt, 2014). The target sound is assimilated into the L1 sounds system, based on similarities and dissimilarities. Mature learners can hear the discrepancies and similarities of phones between the native and non-native languages. They perceived the discrepancies and similarities in terms of acoustic properties or articulatory properties. If the listener perceived the phonemes of the non-native language to be very similar to their native language, then they may not be able to detect the discrepancies. This resulting in the nonnative phoneme being assimilated to the native phoneme category. Contradictorily, if

the listener perceived the phonemes of the non-native language to be different than their native language, then there is a discrepancy detected. Thus, the phoneme of the non-native will not be assimilated to the phoneme of the native language.

This model proposed four categories, 1) two-category, 2) category goodness, 3) single category and non-assimilable category (Best, 1991). This explains how non-native sounds are perceived and assimilated into native sounds. Table 2.4 shows the explanation of each of the categories with examples.

Table 2.4 Perceptual Assimilation Model (PAM) Categories (Komurcu & Yildiz, 2011; Zahariah, 2005)

Category PAM	
Single Category (SC)	Two L2 phones assimilate to one phone in the native category and both are equal to the native sound. For example, American English /æ:/ and /ɑ:/ (L2) are assimilated into a single Japanese phoneme, /a:/. Both phones are perceived as the same.
Two Category (TC)	The phone of L2 contrast is assimilated to two different native phones. For example: English alveolar /s/ and /t/(L2) are assimilated into Persian dental /s/ and /t/ (L1) categories.
Category Goodness (CG)	Each sound of the L2 contrast is assimilated to the same native category, with one of the members is more deviant from the native sound than the other For example: English /ɪ/ and /i:/ (L2) are assimilated into Spanish /i/ but English /i:/ is perceived as a better example of Spanish /i/.
Non-Assimilable (NA)	The L2 phones may be too dissimilar to any L1 phone and cannot be assimilated into any category. For example English phones (L2) to Zulu clicks (L1).

Here are some examples of research that applying PAM into their research as they were looking at the participants' perception of English sounds. Research conducted by Komurcu & Yildiz (2011) specifically looked at the effectiveness of PAM in analysing

L1 transfer. The participants of this research were young adults, age six and four years old. The participants were to identify English minimal pairs, to their native language Turkish (the sounds cannot be identified due to printing error). The findings indicated that PAM's TC and CG explained the assimilation of English to Turkish, in which the young participants have a better perception of the targeted language. In addition, Zahariah (2005) has conducted research on English word-final obstruent by Malay speakers. She studied the perception of /t/ - /d/, /f/ - /v/ and /s/ - /z/ in word-final position. Her findings showed that /s/ - /z/ fell under SC, as they were the most difficult to be distinguished. /t/ - /d/ was the easiest to be distinguished, putting them into the TC. However, /f/ - /v/ result was special, as the result for discrimination task was similar to /t/ - /d/ but participants were struggling during the identification task. There was no clear explanation, and further analysis could be conducted to see if there is a sub-category under TC.

2.4.2 Speech Learning Model (SLM)

The speech Learning Model (SLM) proposed by Flege (1995a, 2007) claimed that the ability to perceive cross-language phonetics will determine the success of producing the L2 sounds (Baker, Trofimovich, Flege, Mack, & Halter, 2008). If the L2 phone is too similar to a phone in the L1, it will be assimilated into the L1 category (refer to PAM Single Category). Thus, the L2 phone will be produced as to how the phone of L1 is produced. However, if the L2 phone is dissimilar to any L1 phone, a new category is created (refer to PAM Two-category). Thus, the L2 phone is produced properly and differently than the L1 phone. What makes SLM is different from PAM is that PAM focuses on the perception of L2 by L1 speakers. However, SLM is looking at both the production and perception of the non-native language by L1 speakers and how they are connected. As mention before, the more similar the two phones are between L2 and L1, the more difficult it would be for the learners to acquire the language. This is because the L2 sound is assimilated into L1, thus causing the learners to perceive the L2 sound as the

L1 sound. This will reduce their awareness of the two languages. On contrary, when the L2 phone is different from the L1 phone, a new category is created. This works well for L1 that has fewer sounds compared to L2, for example, Spanish 5 vowels system vs English 15 vowels system (Flege, 1995a). The learners will perceive these sounds different than the sounds of their L1. With a better perception, the production of the sounds is also improved, as they are specific to their phonetic category representation.

An example of research applying this model is by Ammar et al. (2016). This research was looking at the effect of L1 and proficiency level on English vowel perception by Iraqi speakers. Twelve English monophthongs were included and compared against the native L1, Iraqi Arabic. The findings showed that L1 did influence the acquisition of English and was supported by SLM. The long vowels /i:/ and /u:/ were easily produced as they are present in both languages. Similarly, not much issue was found for /Λ/, /ə/ and /ɜ:/, the central vowel, as these vowels were not present in the L1. Thus, new categories were created to fit them, allowing the speakers to be aware of their existence. Lastly, L2 vowels which were too similar to the L1 posed high difficulties to perceive, such as /ɔ:/, /ɑ:/ and /ɛ/.

2.4.3 Response Latency

Response latency explains that more accurate responses take a shorter time (Schiepers, 1980). In his study, he suggested that there are three stages in word recognition, 1) word concepts activation, 2) decision between possible responses and 3) speech programming, as illustrates in Figure 2.4. The first stage explained that the brain processes and integrates incoming information (from visual especially) and activating appropriate schema in order to understand the concept. In stage two, the mind is making the decision of which response to employ. This is after the activation is complete and the word become available. During this stage, several alternative options may emerge, and the information gathered during

activation will assist in making the correct judgement. The last stage is when the respondent responds to the stimulus, by uttering the word. Thus, response latency is measured during all these stages.

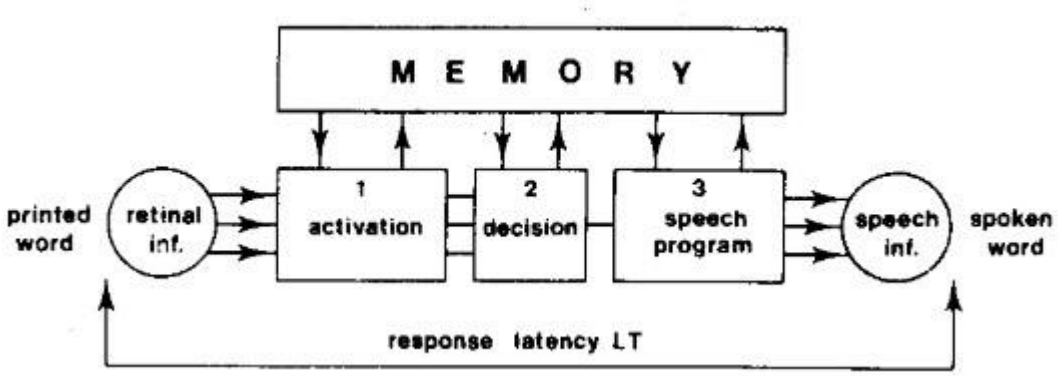

Figure 2.4 Word Recognition Scheme (Adapted from Schiepers, 1980)

The political scientist has been using response latency method in their research, especially on survey. As the participants were interviewed, the time between the question asked and answered were measure. Fazio and William (1986) (as cited in Mulligan, Grant, Mockabee, & Monson (2003)) claimed that accessibility can reflect the strength of the relationship in memory between the object and the evaluation. It is normal for a human to express their honest opinion at the beginning, before changing them to adapt to the situation. Dunning & Perretta (2002) supported this method as their research showed a positive result on latency and accuracy. According to them, "*Accurate positive identifications are more likely to be automatic in nature.*" Although their research was looking at eyewitnesses in identifying the culprit in the line-up, this finding showed how humans, by nature, automatically choose correct/positive responses.

2.4.4 Production-Perception Link

The proper production of a sound is influenced by the ability to perceive it correctly, and this is known as the perception-production link. There is no clear origin of when and who coined this term but it dated back to 1934 (Isbell, 2016). There are a few theories and

models that support this idea such as Fledge's SLM, and Motor Theory of Speech Perception (MTSP). In the perception-production link, perception is important, and it influences the production. Thus, the perception must come first before the production process can take place. Bradlow, Akahane-yamada, Pisoni, & Tohkura (1999) in their work revealed how perception can really improve learner production. They worked with Japanese learners and provided the participants with only perception training for three months. Evaluation by native speakers on the production of English /r/ and /l/ words by the Japanese participants showed improvement in terms of general quality, identifiability, and overall intelligibility. Similar findings were also found in a research by Saito & Poeteren (2018) as they revisit the perception and production of Japanese speakers towards English /r/ and /l/. They concluded that "good perception ability may help L2 learners to achieve more intelligible English /ɹ/ production via the resetting of existing articulatory parameters." Participants with better perception scores produced better target sounds, based on the four dimensions of production ability. In addition, Elvin, Williams, & Escudero (2016) found the relationship of perception and production on L2 and agreed that the perception must precede the production. The learners initially perceived the sound of L2 as how they perceived sound in their L1, causing them to produce L2 sounds as how they produced L1 sounds, as suggested in the Second Language Linguistics Perception Model. Studies such as by Kendall & Fridland (2012) and Kleber & Harrington (2011) found a clear relationship between perception and production.

2.5 Conclusion

Malay and English have a similar vowel system, but of course, they are not the same. Due to this similarity, Malay speakers of English encountered a problem to properly produce the L2 sounds accurately. Realising this issue, the Malaysian education system is continuously undergoing improvement in hope of improving the ability to communicate in English. Previous researches showed that the issue of perceiving and producing English

sounds, focusing on vowel sounds, is common among second language learners. Those researches, however, looking at adult and young adult speakers. Moreover, much Malaysian research only looked at either production or perception of the English language alone, without making a connection between the two. Thus, the present research is looking at the production and perception of English vowels by Malay young speakers.

CHAPTER 3: RESEARCH METHODOLOGY

This chapter describes the methodology employed in this research. This chapter has five sections. The first section explains the participants who were chosen and their criteria for being selected. The second section presents the chronology of acquiring consent from all parties. After that, section 3 describes the method of data collection. In the last two sections, section 4 talks about the procedure and finally, section 5 explains the data analysis for this research.

3.1 Participants

The participants for this research were 10 primary school female pupils from Sekolah Kebangsaan Desa Setapak, Kuala Lumpur. Participants were selected based on specific criteria; 1) they are Malay ethnicity, 2) English is their second language, 3) they are Year 5 pupils and lastly 4) their gender is female. The reason for criteria (1) is that this research is looking at the ability of Malay young speakers in producing and perceiving English vowel sounds. Thus, it is important to select Malay speakers who speak Malay as their native language. To ensure criteria (2) is fulfilled, Sekolah Kebangsaan Desa Setapak was chosen. The school is located in the suburban and the population is middle-class families. This research is looking for participants who did not speak English in their daily life. Middle-class families that live in the suburban area are adequately being exposed to the English language, but the usage is limited. English. English is used in a formal setting such as workplace and school, but not in their household. Another important element in this research is the ability to use a computer by the participants. This research employs computer applications, thus a basic understanding of handling a computer program will help the data collection process. Rural area students may not have much experience with computers, and it may cause the researcher to spend more time explaining and briefing.

This could increase the participants' anxiety, which may affect the end product of the data collected. Moreover, parents and guardians are more willing to cooperate as they understand the importance of this research. Year 5 pupils (criteria 3) were selected due to their maturity in understanding the purpose of the research and their four years' experience of studying English formally in school. Finally, criteria (4) was chosen as only female students were selected is to ensure the gender variable is constant (Burman, Bitan, & Booth, 2008; Schepens, van der Slik, & van Hout, 2015). Males' vocal tract is wider than females' vocal tract, and this would influence the production of the sound frequency of the vowels, thus may affect the results. For this research, only 10 of them were selected due to controlling the number of data collected for the analysis.

3.2 Consent

This research involves primary school and primary school pupils. Thus, permission from the Ministry of Education Malaysia is sought. A prior meeting was conducted between the researcher and the headmaster of Sekolah Kebangsaan Desa Setapak to acquire permission on conducting the research. After all, the recording and test will be conducted in the school compound, during school hours. After the request was granted, a formal request was sent to Bahagian Perancangan dan Penyelidikan Dasar Pendidikan (BPPDP), Kementerian Pendidikan Malaysia and the permission was acquired (refer to Appendix A). Letter of Consent was given to the parents of the participants through the school administration (see Appendix B) due to the participants are underage.

3.3 Data Collection

This research employed three methods of data collection. The three types of instruments used to collect the data are language profile test, production test and perception test.

3.3.1 Language Profile Test

The Language Profile Test (refer to Appendix C) was used to assist the selection of the participants. The test was adapted loosely from (Birdsong, Gertken, & Amengual, 2012). The test was prepared in Malay, to ease the participants' understanding and answering process. The test consists of two parts; Part 1 is Demographic details such as name, age and gender, and Part 2 is the Language Profile Test. There are 12 questions in Part 2; 1 question requires participants to write their answer, 4 questions require them to circle the best answer and 7 questions of the Likert Scale option. This test was used to identify the participants, as the questions in the test were constructed based on the criteria set earlier (refer 3.1). The demographic part identified the gender, race and age of the participants. Part 2 of the test looked into the language profile, filtering either the potential participants were using Malay as their native language and English as their second language. Moreover, the self-rated English fluency was to check the participants' language confidence, as well as the ability to produce and perceive English in general.

3.3.2 Production Test

This research focused on English short and long vowel sounds; hence three sets of vowel pairs were selected; /ɪ /- /i:/, /ʊ/ - /u:/, and /ʌ/ - /ɑ:/. The vowel sounds were embedded in a CVC context word, where C is a plosive consonant; /b/, /p/, /d/, /t/, /k/ and /g/, and placed in a carrier sentence: "The word is CVC". Each vowel has three words representing them, two words following the rule above and one acted as a distractor. This was to ensure the production of each word (especially the targeted sound) was done as naturally as possible. Table 3.1 shows the list of the words for the production test. Condition 'TW' referred to targeted words and 'D' referred to distractor words.

Table 3.1 Word List for Production Test

	Short Vowel	Long Vowel	Condition

1.	bid	/bɪd/	beep	/bi:p/	TW
2.	mitt	/mɪt/	meat	/mi:t/	D
3.	pick	/pɪk/	peak	/pi:k/	TW
4.	book	/bʊk/	boot	/bu:t/	TW
5.	cook	/kʊk/	coop	/ku:p/	TW
6.	foot	/fʊt/	food	/fu:d/	D
7.	cut	/kΛt/	cart	/kɑ:t/	TW
8.	duck	/dΛk/	dark	/dɑ:k/	TW
9.	hut	/hΛt/	heart	/hɑ:t/	D

3.3.3 Perception Test

This experiment is adapted loosely from (Casillas, 2015). 2 Alternative Forced Choice (2AFC) was used rather than the typical ABX method because aside from accuracy, this research also looked at the latency. Reaction time requires stimulus perception, decisionmaking process and execution of response (Ng & Chan, 2012). Providing more than two choices will affect the 'thinking time', thus increasing the reaction time. Other than that, this psychophysical method is unbiased and more accurate. Researchers such as Flege

(1995b), Kondaurova & Francis (2008), Saito & Poeteren (2018), Schiepers (1980) and Zahariah (2005) also applied this method in their research.

The words list was designed based on these criteria, 1) the words chosen must be among common objects, to ensure the participants are able to recognise them, 2) the object must be a concrete object and lastly the words must contain the target vowels. **Table 3.2 Word List for Perception Test**

	Short Vowel	Long Vowel
1.	Bin	Bean
2.	Book	Boot
3.	Hut	Heart

Table 3.2 shows the list of words used in the perception test. The selected words followed the CVC context (similar to 3.3.2) and images of the words, specifically cartoon images were used during the test rather than the words (spelling) and this is employed for a better interpretation (refer to Appendix D). According to Casillas & Simonet (2016),

images are used to reduce confusion as well as to reduce orthographic cues in inferring answers and this is supported by other researches such as Escudero & Boersma (2004) and Kondaurova & Francis (2008). The word was then placed in a carrier sentence filling the blank space,

_____, the word is _____

The stimuli, reading for the sentences were prepared and the reading was recorded by a Malay speaker of English, who was very proficient in the English language (using the same setup as the production experiment). The speaker has a doctoral degree in English studies and has been working in English education for more than 20 years. The speaker was selected in order to reduce accent interference as well as ensuring the sounds produced are as familiar as possible to the participants (Casillas & Simonet, 2016). The F1, F2 and duration of the targeted vowels were measured, using Praat software, and showed in Table 3.3. The duration differences were obvious for all three pairs: 106.62 msec for /ɪ/ and /i:/ pair, 136.94 msec for /ʊ/ and /u:/ pair and finally 206.24 msec for /ʌ/ and /ɑ:/ pair. Three English speakers have rated the recordings, who work in the English language field and specialising in phonetics and phonology, for perceptual verification. The rating indicated that the audio did illustrate differences between each pair word, hence perceptually reliable and valid.

Table 3.3 The value of duration (msec), F1 and F2 (Bark) of the stimuli

Vowel	Duration (msec)		F1 (Bark)	F2 (Bark)
	Actual	Difference		
/ɪ/	54.75	106.62	3.739	13.645
/i:/	161.36		3.228	14.611
/ʊ/	93.56	136.94	4.207	8.795
/u:/	230.50		3.884	10.298
/ʌ/	60.40	206.24	7.818	11.296
/ɑ:/	266.64		7.836	9.796

The software Pyschopy version 3.0.0b6 (Peirce, 2017) was used to present the stimuli (reading of the sentence) to the participants (Peirce, 2008). The task consisted of four parts; Part 1 is for exercise; which to help participants to be familiar with the software, Part 2 presents the /ɪ/ - /i:/ pair, Part 3 presents the /ʌ/ - /ɑ:/ pair, and Part 4 presents the/ʊ/ - /u:/ pair.

3.4 Procedure

The process of data collection started with getting permission from authorities in order to conduct this research. A meeting with the headmaster of Sekolah Kebangsaan Desa Setapak was held and permission was granted. Since the national school falls under Ministry of Education jurisdiction, hence a formal request to perform the research in the school compound as well as to have the primary school pupils as the subject of this research was sent to BPPDP. Once the permission was granted, a Letter of Consent was distributed to all parents and guardians of the potential participants, with the assistant of the school management. This process took approximately a month to complete, and after completing this step, the research in then commence.

Before the actual research was conducted, the wordlists (Table 3.1 and Table 3.2) were shown to the class's teacher to ensure all words chosen are familiar to the pupils. Two weeks after completing the Language Profile Test, the production and perception tests were then conducted. Both tests were done within the same day, within the school compound during school hours. The school has provided two computer labs for the researcher disposal; thus, Computer Lab 1 was used for the production test while Computer Labs 2 was prepared for the perception test. At the same time, the school allocates two teachers to assist the researcher.

3.4.1 Selection of Participants

There were 26 Year 5 pupils taking part in this test, both male and female pupils. A slot for them to sit for the Language Profile Test was set with their class teacher, allowing everyone that present to undergo the test. This test was held during school hours, in their own classroom. Once the test was answered and collected, data from the test were analysed. The test was conducted two weeks prior to the actual research; the production test and perception test.

The results from the Language Profile Test were used to identified and filtered possible participants. In Part 1, 17 female pupils were identified, and the male pupils were then excluded from the research. Part 2 of the test provided more information to further filter and select the participants. Here, it was found that all participants acquired Malay as their first language, as that is the language they have been exposed to since birth. As for English, a few participants had some exposure to the language as early as 3 years old, but most of them learn the language at the age of 7 (or Year 1). As expected, all participants preferred to communicate (with family or in class) using Malay and they are more comfortable speaking in Malay than in English. The self-rated English fluency showed participants' language confidence, as well as the ability to produce and perceive English in general. Out of 17 potential participants, only ten were finalised and selected, five pupils' parents did not agree to the involvement and the other two refused to participate themselves.

3.4.2 Production Test

On the day of the research, the researcher and the two teachers gathered the participants outside the computer lab. The participants then were briefed on the purpose of the research as well as the procedure of the two tests. The research started with the production test first then followed by the perception test. The participant was called one by one; hence the rest

of the participants were waiting at a nearby gazebo. Once they were called, they will go into Computer Lab 1 for the production test, and later they will go to Computer Lab 2 for the perception test. After completing both tests, the participants will wait for the research to finish at a nearby gazebo.

A Sony ICD-UX560F digital recorder, set at 44.1kHz and 16-bit quantization and a Takstar SGC 578 Shotgun microphone were used to record the reading session. The microphone was set on the participant's right, slightly higher than her eye level to avoid capturing breathing sounds. It was also to ensure the vocal track is free from blockage as the participant was required to face up during the reading session. The stimuli were presented to the participant using the PowerPoint slides presentation. Each sentence was presented in a slide, to avoid confusion as well as provide enough time for the participants to recite the sentence completely. Once a sentence has been recited, the next slide was presented by clicking the next button. Digital recorder was used as the data can be easily handled and transfer to computer, later to be uploaded into Praat software, as suggested in Casillas (2015) and Jacewicz, Fox, & Salmons (2011). The recordings were saved in the WAV format, uncompressed audio format as it enables better manipulating and simplifying data.

The test started by allowing the participant to sit comfortably facing a computer. The computer displayed the sentence list for the participant to read out loud. The participant was given five minutes to read through the text first. Additional five random sentences were displayed, and the participant was asked to read them. This acted as an ice breaker and helped the participant felt more comfortable. This reading was not recorded. As the participant reached the sixth sentence, the researcher started the recording of the reading. Each target sentence was read three times. Thus, 54 tokens were produced by the participants. Each of the vowels is represented by three words that were embedded in a

sentence. This means each vowel has three sentences representing them, and each sentence is read three times (6 vowels x 3 sentences x 3 repetitions). However, only targeted sentences will be analysed and distractor sentences will be excluded. Hence, only 36 tokens per participant were selected and analysed (6 vowels x 2 sentences x 3 repetitions). In total, 360 tokens were recorded from all 10 participants to be measured and analysed (6 vowels x 2 sentences x 3 repetitions x 10 participants).

3.4.3 Perception Test

The perception test is presented to the participants using an HP ProBook 4430s laptop to run the Pyschopy software and a Logitech headphone was prepared for the listening part. The task consisted of four parts; Part 1 is for exercise, Part 2 is the /ɪ/ - /iː/ pair, Part 3 is the /ʌ/ - /ɑː/ pair, and Part 4 is for /ʊ/ - /uː/ pair. The exercise part is for the participants to get familiar with the experiment, especially on how to use the software. This is necessary as Racine (2013) in his research discussed how reaction time is sensitive; having a practice before the actual experiment can prevent any undue influence. When each part started, the screen showed the carrier sentence, "_____, the word is _____" and the recording of the targeted word was played. Simultaneously, two images of a possible answer appeared at the bottom part of the screen. The screen will only continue to the next reading once the participant answered by pressing the '**LEFT**' or '**RIGHT**' button on the keyboard. Five seconds break was given between each part to allow some recovery time. Figure 3.1 shows the sample of the Psychopy software interface.

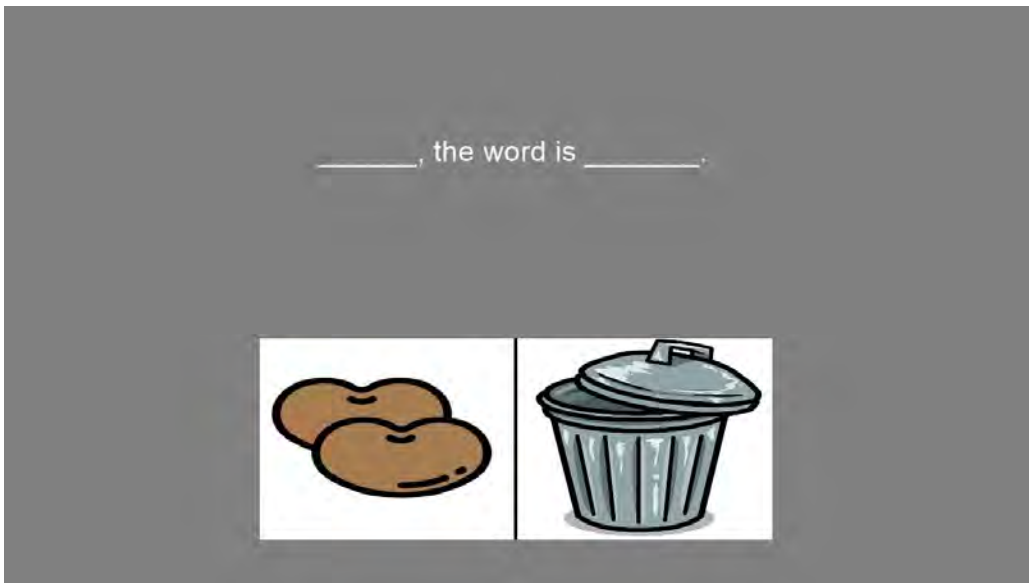

Figure 3.1 Screenshot of the interface of Psychopy software

Once the participant completed the production test, she moved to Computer Lab 2 for the perception test. The participant was seated comfortably facing the laptop. The teacher helping the researcher was there to give a briefing on how to use the software and to assist any inquiry from the participant. Once ready, the participant wore the headphone and started the task. After completed the task, the participant was required to inform the teacher and exited the lab. This process continued until all 10 participants completed their turn. Each vowel will have three repetitions, and there are six targeted vowels. Thus, each participant will produce 18 responses or tokens (6 vowels x 3 repetitions). In total, 180 tokens were produced by all 10 participants (6 vowels x 3 repetitions x 10 participants).

A set of brooches was given to each participant as an honorarium from the researcher. The experiments took about half an hour to complete with each participant took not more than 10 minutes to complete both experiments. The data collected is then measured, tabulated and analysed.

3.5 Data Analysis

The data collected earlier is then measured, tabulated and analysed. The analysis was done based on the instruments and appropriate methods.

3.5.1 Language Profile Test

The Language Profile Test is used to understand the participants' language background and fluency. Table 3.4 shows the selected participants' language background and their self-rated English fluency. From the table, the earliest time the participant received exposure to English was at the age of three (P1 & P6), while a majority of them received exposure to the English language during their formal education session, seven years old or equivalent to Year 1 (P2, P3, P4, and P5). When asked to rate their English fluency, only P1 rated herself to be fluent in speaking, while the majority chose not very fluent. Similarly, P1 rated herself to be good in listening too, while most of them rated themselves with '2'. The result surprisingly showed that the participants rated their English reading and writing lower compared to listening and speaking.

Table 3.4 Participants' Language Background

Participant	Exposure to Malay	Exposure to English	Language most comfortable	English speaking fluency	English listening fluency	English reading fluency	English writing fluency
P1	At birth	3	Malay	4	3	3	2
P2	At birth	7	Malay	2	1	1	1
P3	At birth	7	Malay	1	1	2	1
P4	At birth	7	Malay	1	1	2	1
P5	At birth	7	Malay	1	1	1	1
P6	At birth	3	Malay	2	2	1	2
P7	At birth	4	Malay	2	2	1	1
P8	At birth	6	Malay	1	2	1	1
P9	At birth	6	Malay	2	2	1	1
P10	At birth	5	Malay	2	2	1	1

3.5.2 Production Test

The sound files were transferred to a laptop, labelled accordingly and analysed using PRAAT version 6.0.37 (Boersma & Weenink, 2017). The file was labelled according to the participant's name and number (e.g: Alia_1), hence each participant has three sound files. Using the TextGrid function in Praat software, the data was transcribed, isolating the target vowel to be measured and analysed. The targeted vowel was isolated from the rest of the sentence using the Tier function.

1. Tier 1 was the whole sentence.
2. Tier 2 was the targeted word.
3. Tier 3 was the targeted vowel (labelled with the phonetic symbol of the vowel).
4. Tier 4 was the F1 value.
5. Tier 5 was the F2 value.

The onset of a vowel was identified at the beginning of a formant structure after the release of initial C, and the offset by the point before the next C's cessation of the acoustic signal, marking the boundaries for the targeted sound and measuring the duration (refer to Figure 3.2). The F1 and F2 values were measured at the midpoint of a vowel, where a vowel is most stable and reliable to measure (Casillas, 2015; Hawkins & Midgley, 2005; Hubais & Pillai, 2010; Perwitasari et al., 2016). This was done by running the formant script in PRAAT, automatically measured based on the linear predictive coding (LPC) in the software. The duration of the vowel sound was measured from the onset to the offset of the vowel, in millisecond (msec), by running the duration script in the software.

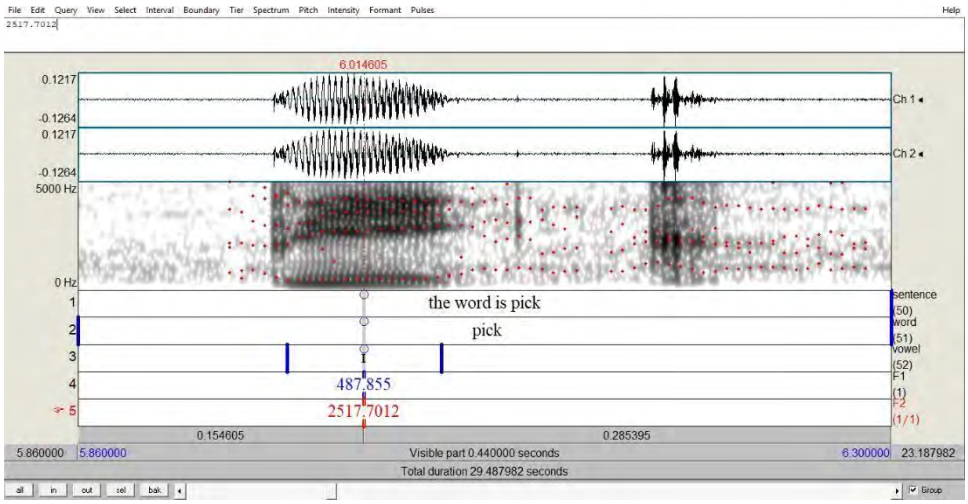

Figure 3.2 Screenshot of the TextGrid in Praat software for the work 'pick'

Then, the data were transferred to an Excel sheet and graphs were plotted. The value of F1 and F2 were converted to Bark in order to normalise the value (Yamaguchi & Chiew, 2019). This is because Bark distance corresponds with the perceptual distance and a good approximation of the actual frequency analysis performed by the ear (Kent & Read, 2002; Saito & Poeteren, 2018). The formula used is as followed, proposed by Traunmüller (as cited in Thomas (2010)).

$$\text{Bark} = [26.81/(1+1960/\text{Hz})]-0.53$$

The Euclidean Distance (ED) is measured to understand and predict the possibilities of discrimination among the vowels produced. The formula used in this research is as follows:

$$d = \sqrt{(F1i - F1j)2 + (F2i - F2j)2}$$

d = distance between the variables i = first vowel under investigation j = second vowel under investigation F1i = first formant frequency value for vowel
F1j = first formant frequency value for vowel
F2i = second formant frequency value for vowel
F2j = second formant frequency value for vowel

3.5.3 Perception Test

Data of each participant's response from Psychopy was imported and combined into an Excel sheet. The data was then filtered accordingly, unimportant information such as distractor trials was removed, and variables were renamed to a clearer version. The important variables in this data are:

1. Sound File (or the word used).
2. Vowel types (refers to short or long vowels).
3. The vowel pairs.
4. Participants' responses.
5. Participants' reaction time.

The accuracy of the responses was calculated, the correct answer represented by '**1**' and the wrong answer was indicated by '**0**'. The latency was recorded in seconds. Data were then tabulated and charts were produced.

3.5.4 Statistical Analysis

The data from the production test and perception test were then analysed following the statistical analysis using SPSS (IBM Corp, 2010). A paired-samples T-Test was used in analysing the data as the two data compared are gathered from the same participants. This test is used to determine the significant differences between the vowel pairs; /ɪ/ - /i:/, /ʌ/ - /ɑ:/, and /ʊ/ - /u:/. The significant differences between English short and long vowels were identified in the form of frequency values (F1 and F2) and the durations in the production test, as well as the accuracy and latency of the vowel pairs in the perception test. This research hypothesis is that there is a clear difference in the value of F1, F2 and duration between the short and long vowel in the production of the sounds, while in perception, the participants can differentiate between short and long vowels by achieving

the highest accuracy in identification test in the shortest time. The significant level is set at the common value of 0.05.

3.6 Conclusion

This chapter describes the methodology used to examine the ability of Malay young female speakers in producing and perceiving English short and long vowels. Thus, a Profile Language Test was used to identify the participants. Then, the production test was used to identify the vowel quality and vowel duration, while the perception test was used to test accuracy and latency in distinguishing the monophthong pairs. All data are tabulated and analysed, charted into tables and graphs. The following chapter will present and discuss the results of the research.

CHAPTER 4: FINDINGS AND DISCUSSION

This chapter has two sections that present the findings (section 1) and the discussion (section 2) of the study. Section 1 is divided into two parts: 1) Production Test and 2) Perception Test. The first part presents the result of the vowel productions produced by the participants. Vowel charts were plotted and graphs were tabulated to illustrates the F1 and F2, as well as the duration of the vowel, studied better and clearer. As for the second part, the results of the listening test are illustrated and analysed in graphs and the findings of the accuracy and latency are presented. After that in Section 2, the findings are discussed in three parts: 1) discussion on the production test, 2) discussion on the perception test and 3) the relationship between production and perception test.

4.1 Production Test

The production test analysed the sounds produced by the participants. The targeted vowel sounds were embedded into carrier sentences for the participants to read and the readings were recorded. The recording was then evaluated, targeted sounds were isolated and tabulated. The results of the production are presented by looking at the vowel quality and vowel duration.

4.1.1 Vowel Quality Contrast

As discussed in the previous chapter, the F1 and F2 represent the position of the vowel in the vowel chart. The frequency of F1 indicates the vowel height (y-axis) while the frequency of F2 refers to the front and back of the vowels (x-axis). As the frequency of F1 increases, it shows that the tongue is lowering and the jaw is being opened, whereas the increase of F2 value corresponds with the tongue move towards the front of the mouth. The values of F1 and F2 were then charted to understand the quality contrast (refer to Figure 2.2 for the standard English vowel chart). In this research, The F1 and F2 values

of the participants were measured and tabulated. The data was then converted into charts as follows.

4.1.1.1 /ɪ/ and /i:/ Pair

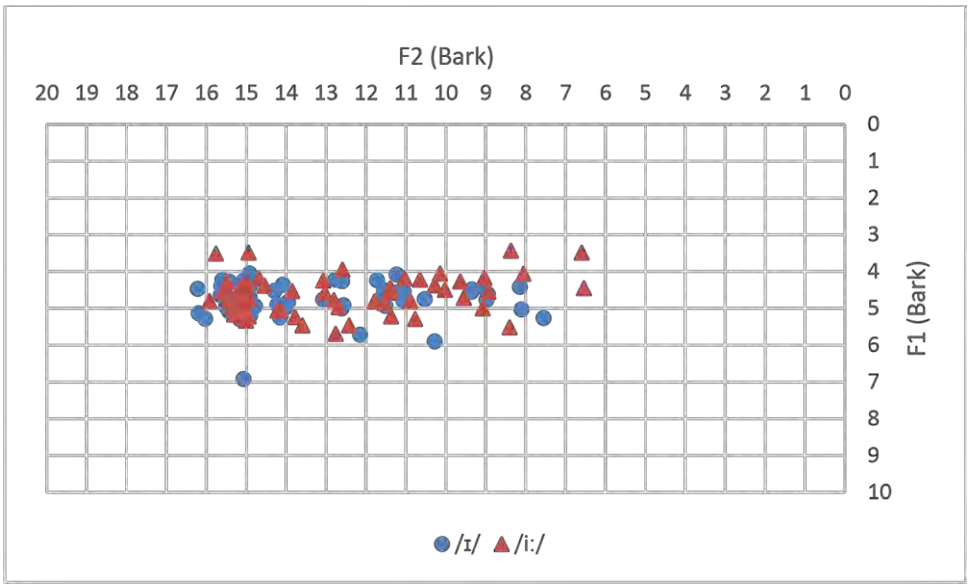

Figure 4.1Vowel Distribution for /ɪ/ and /i:/

Figure 4.1 shows the distribution of the vowels/ɪ/ and /i:/ produced by the participants, as they were plotted into vowel charts. The average value of F1 for /ɪ/ is 4.733 Bark (SD = 0.480 Bark) and F2 at 13.278 Bark (SD = 2.437 Bark), while the average values of F1 for /i:/ is 4.621 Bark (SD = 0.509 Bark) and F2 average value at 12.757 Bark (SD = 2.610 Bark). The values of F1 for /ɪ/ are distributed between 4 to 7 Bark. The plotting concentrates mostly at 4 and 5 Bark, and less at 7 Bark with only one plotting. As for the value of F2, the plottings are widely distributed from 7 to 16 Bark. The chart shows two major clusters of concentration plotting, cluster 14 to 16 Bark has the highest plotting, follows by clusters 11 and 12 Bark. The numbers of plotting decrease with the value of F2 decreases. A similar result can be seen with the values of F2 for /i:/. The values are widely distributed along 6 to 16 Bark, with the majority of the plotting can be found around 15 and 16 Bark. The numbers of plotting decrease as the Bark value decreases. As for the F1 value, the values are distributed from 3 to 6 Bark, and the concentrates are mostly around 4 and 5 Bark.

From the figure, the vowel realization of \ɪ\ and \i:\ were not being clearly distinguished. The plotting is mostly on the front, consistent with the position of the pair in the chart. The plotting is mostly overlapping, with \ɪ\ distribution is lower than \i:\. However, \i:\ plotting is more back compares to \ɪ\ in the chart, which indicates these two vowels have swapped places.

4.1.1.2 /ʌ/ and /ɑ:/ pair

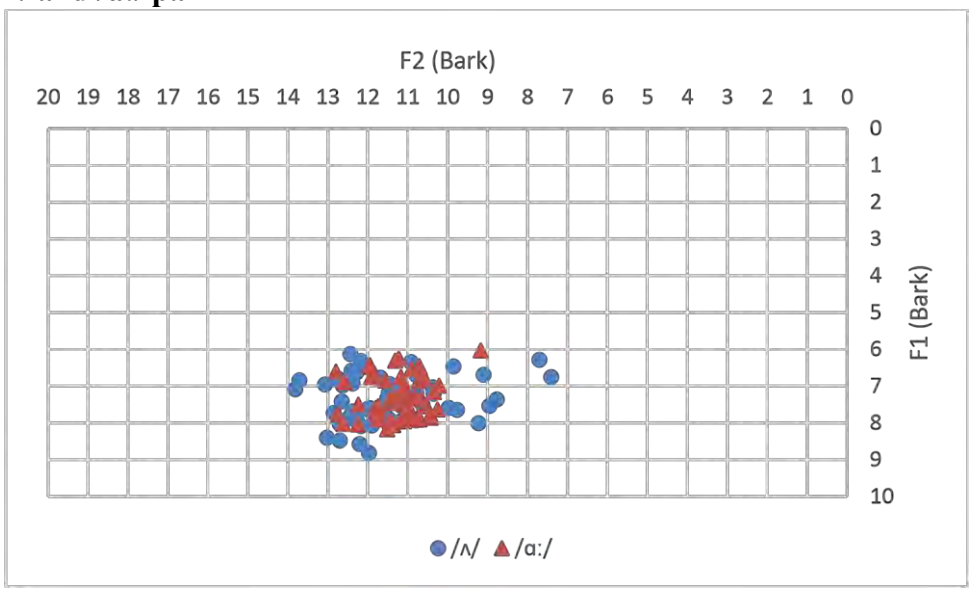

Figure 4.2 Vowel Distribution for /ʌ/ - /ɑ:/

Figure 4.2 shows the plotting of vowel distribution for /ʌ/ and /ɑ:/. The average value of F1 for /ʌ/ is 7.373 Bark (SD = 0.614 Bark) and F2 at 11.436 Bark (SD = 1.319 Bark), while the average values of F1 for /ɑ:/ is 7.339 Bark (SD = 0.553 Bark) and F2 average value at 11.238 Bark (SD = 0.673 Bark). The distribution of F1 values of /ʌ/ is towards the bottom part of the chart, covers between 6 to 9 Bark. The plotting mostly occurs at 7 and 8 Bark. As for the F2 values, wider distribution is plotted, with values from 7 to 14 Bark. On contrary, the plotting of /ɑ:/ is more compact, with the F1 values are between 6 and 8 Bark whiles the F2 values are between 9 to 13 Bark. This shows that there is a higher possibility that the participants can produce the sound better than its counterpart.

The plotting of both /ʌ/ and /ɑ:/ are mostly low and almost centre, consistent with the position of both sounds in the vowel chart. Comparing their positions, /ʌ/ is produced more to the front than /ɑ:/, but it is difficult to clearly identify which vowel is higher in the chart. A better result can be seen in the plotting of the average values of the Bark.

4.1.1.3 /ʊ/ and /u:/ Pair

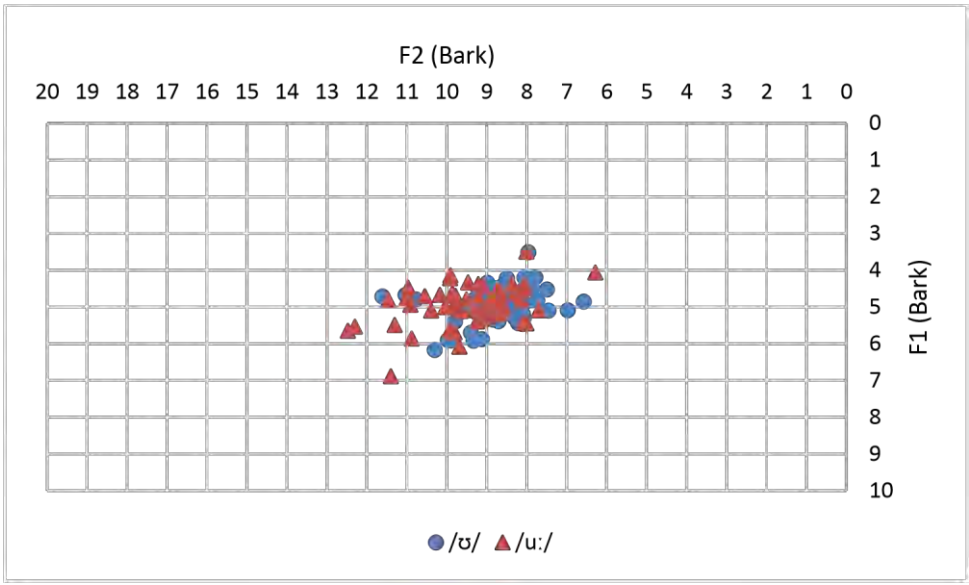

Figure 4.3 Vowel Distribution for /ʊ/ and /u:/

Figure 4.3 shows the distribution of vowels/ʊ/ and /u:/ by the participants in the vowel chart. The average value of F1 for /ʊ/ is 4.934 Bark (SD = 0.483 Bark) and F2 at 8.686 Bark (SD = 0.894 Bark), while the average values of F1 for /u:/ is 4.947 Bark (SD = 0.528 Bark) and F2 average value at 9.376 Bark (SD = 1.154 Bark). Similar to /ɑ:/, the production of /ʊ/ is more compact and focused. The F1 values are distributed between 3 to 6 Bark and the F2 values are distributed between 6 to 12 Bark. The values are mostly produced at 4 and 5 Bark for F1 value and 8 to 9 Bark for F2 value. However, the values of F1 for /u:/ are produced between 3 to 7 Bark, a bit wider compares to /ʊ/. Correspondingly, the distribution of F2 values of /u:/ is also slightly wider to /ʊ/, with values between 6 to 13 Bark.

The /ʊ/ distribution is closer compared to /u:/. Their position is more centre-back and middle-top, almost consistent with the position of this pair in the English vowel chart. When putting them together, /ʊ/ is more back than /u:/, and again it is difficult to clearly clarify their high/low position. A better result can be seen in the plotting of the average values of the Bark.

4.1.1.4 Analysis of the Findings

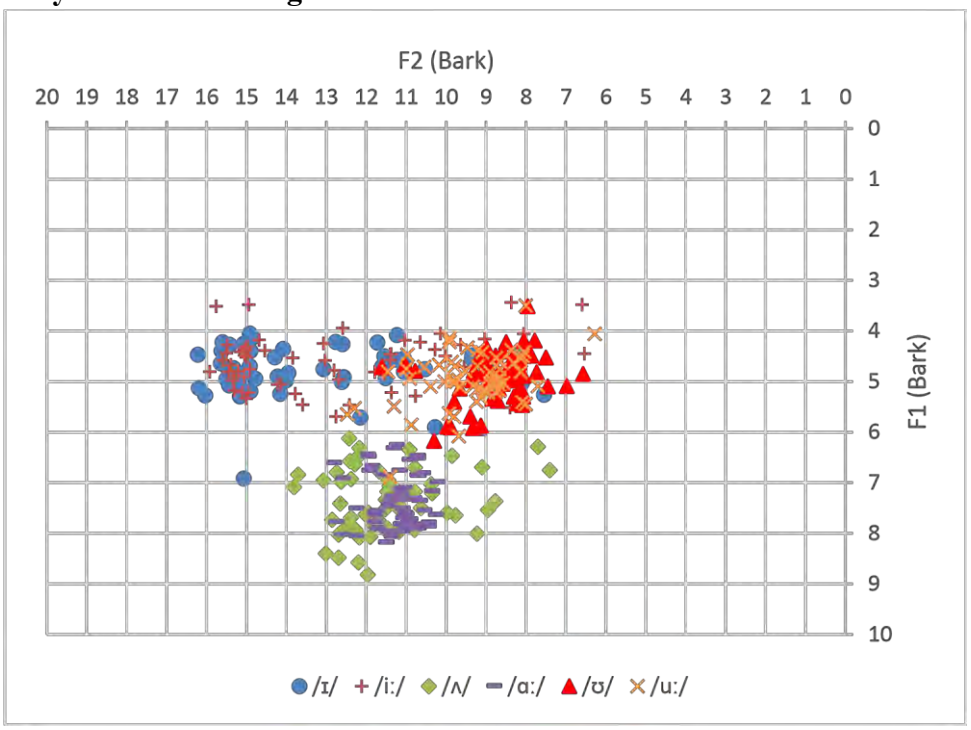

Figure 4.4 Participants' Vowel Distribution

Figure 4.4 shows the distribution of targeted vowels by the participants. The distributions are closer and smaller. The vowel pairs are mostly overlapping. The distributions of /ɪ/ and /i:/ are inconsistence, even crosses the /ʊ/ and /u:/ pair's area. The height-ness of /ɪ/ and /i:/ and /ʊ/ and /u:/ pairs are almost similar, indicates the participants manage to produce the F1 properly. As for the /Λ/ and /ɑ:/, no other sounds overlap with them. Fortunately, the /ɪ/ and /i:/ pair is position at the top-front, /ʊ/ and /u:/ pair is at the top-back and /Λ/ and /ɑ:/, is at low-middle; are similar to their position in English vowel chart. The further analysis looks for the average values of the F1 and F2 frequency of each vowel and the data are tabulated in Table 1.

Table 4.1 Mean Value for F1 and F2 of Targeted Vowel in Hz and Bark, Euclidean Distance in Bark, and Standard Deviation (in parenthesis)

Target Vowel	F1 (Hz)	F2 (Hz)	F1 (Bark)	F2 (Bark)	Euclidean Distance (Bark)
/ɪ/	408 (56.872)	2210 (703.141)	4.733 (0.480)	13.278 (2.437)	2.19
/iː/	467 (56.825)	2060 (715.356)	4.621 (0.509)	12.757 (2.610)	2.14
/ʌ/	822 (90.662)	1606 (297.163)	7.373 (0.614)	11.435 (1.319)	1.52
/ɑː/	817 (79.788)	1540 (158.681)	7.339 (0.553)	11.238 (0.673)	1.54
/ʊ/	503 (56.155)	1035 (160.942)	4.934 (0.483)	8.686 (0.894)	2.36
/uː/	505 (62.127)	1164 (224.005)	4.947 (0.528)	9.376 (1.154)	1.89

Table 4.1 shows the mean values of the F1 and F2 of the targeted vowels for both Hz and Bark, as well as Euclidean Distance (Bark) and the Standard Deviation (in parenthesis) produced by the participants. The data was collected in frequency (Hz), however for analysis purposes and plotting the vowel chart, the data was then converted into Bark. According to the standard English vowel chart, the position of /iː/ is on the top left area, higher and more front compare to /ɪ/. However, F1 and F2 of /ɪ/ (4.733 Bark and 13.278 Bark) are higher than /iː/ (4.621 Bark and 12.757 Bark). This indicates that the position of /ɪ/ is more front than /iː/, as these two sounds have swapped F2 places in the chart but maintain the F1 locations. /Λ/ has more in value for both F1 and F2 (7.373 Bark and 11.435 Bark) than /ɑː/'s (7.339 Bark and 11.238 Bark), thus making the vowel to be more front and lower compared to its long vowel. Lastly, /uː/ (4.947 Bark) and (9.376 Bark) has more values of F1 and F2 compared to /ʊ/, leading this vowel to be positioned more front and lower in comparison to /ʊ/ (4.934 Bark and 8.686 Bark).

In addition, Figure 4.5 illustrates the position of the average values of the targeted vowels (in Bark), plotting the vowel chart for the participants. In general, the vowel qualities produced were almost similar, between F1 and F2 of each vowel pair. Further

analysis shows that /ɪ/ and /i:/ were produced quite close to each other. Similarly, /Λ/ - /ɑ:/ were also very close, with /ɑ:/ was slight to the back. The vowel pair /ʊ/ - /u:/ has the furthest distance of F2, with the F1 was almost similar. A paired-sample t-test was conducted to compare the difference in F1 and F2 of all three short and long vowel pairs. Vowels /ɪ/ and /i:/ produce no significant different in F1 (t (59) = 1.802, p= .077) and F2 (t (59) = 1.322, p = .191). Similarly, vowels /Λ/ and /ɑ:/ produce no significant different in F1 (t (59) = .614, p=.541) and F2 (t (59) = 1.099, p = .276). However, /ʊ/ and /u:/ show no significant different in F1 (t (59) = -.204, p=.839), but there is a significant different in F2 (t (59) = -3.967, p = .000).

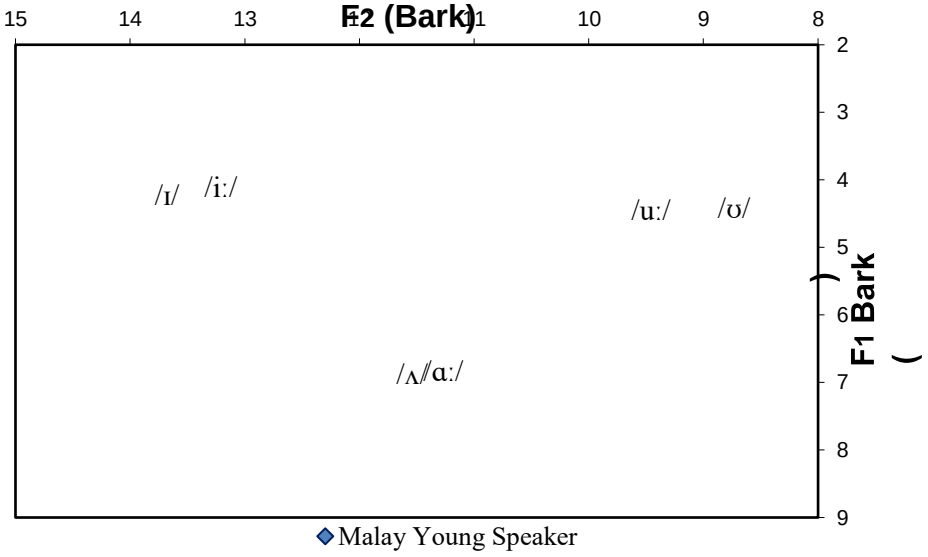

Figure 4.5 Vowel Plotting of the Female Young Learners

The Euclidean Distance measures the distance between each vowel to the centroid of the plotting. /ʊ/ has the longest distance from the centroid of 2.35 Bark, follows by /ɪ/ with the value of 2.19 Bark. /i:/ distance is 2.14 Bark and /u:/ is measured to be 1.89 Bark away from the centroid. /ʌ/ has the shortest distance with 1.52 Bark, follows by /ɑ:/ with 1.54 Bark. These data confirm that the vowel space area produced by the participants is smaller and compact.

4.1.2 Vowel Duration Contrast

English is a sound-based language and the length of the vowel plays an important role. English monophthong consists of short and long vowel pairs, where the short vowel is produced shorter than the long vowel. However, the distinction of the difference of length between the vowel pair is minuscule, sometimes rarely realised. Thus, this research will look at how short is a short vowel and how long is long vowel produced by the participants. Using the same set of data in vowel quality, the duration of the vowel is measured (in milliseconds). The data is tabulated, then converted into charts for further analysis.

Table 4.2 Duration of the English Monophthong Vowel by Female Malay Young Speakers (in msec) and Standard Deviation (in parenthesis)

Participant	/ɪ/	/i:/	/ʊ/	/u:/	/ʌ/	/ɑ:/
P1	134.52 (29.88)	144.62 (30.63)	129.64 (43.26)	139.88 (49.98)	128.10 (17.97)	171.59 (35.74)
P2	88.59 (9.16)	88.59 (22.88)	120.70 (18.96)	120.40 (21.85)	120.20 (21.09)	200.26 (53.32)
P3	102.45 (15.49)	98.32 (16.46)	105.62 (15.50)	104.51 (12.08)	108.88 (20.74)	109.67 (12.40)
P4	95.49 (26.01)	83.20 (11.09)	100.92 (4.51)	92.80 (16.60)	108.82 (11.08)	133.98 (19.47)
P5	66.62 (5.70)	68.48 (15.49)	73.43 (9.98)	65.10 (14.96)	66.19 (9.12)	97.20 (7.39)
P6	100.93 (33.97)	124.12 (47.22)	94.03 (8.93)	101.87 (43.79)	107.55 (17.06)	147.80 (44.76)
P7	112.76 (10.69)	131.22 (27.37)	116.50 (9.96)	124.91 (43.09)	90.89 (10.83)	153.58 (39.26)
P8	104.01 (28.91)	103.98 (16.26)	122.82 (10.23)	141.32 (40.30)	110.16 (41.87)	190.92 (24.87)
P9	111.92 (12.30)	140.92 (32.94)	132.16 (22.19)	130.11 (35.18)	136.83 (16.57)	204.69 (31.69)
P10	118.14 (30.44)	95.27 (14.79)	85.78 (19.00)	92.72 (27.68)	104.63 (8.53)	162.58 (19.14)
Mean	103.54 (28.69)	107.87 (35.94)	108.16 (27.01)	111.36 (40.74)	108.23 (27.25)	157.23 (47.54)

Mean difference between the long and short vowel	4.33	3.20	49.00

Table 4.2 shows the participants' vowel duration in millisecond (msec), Standard Deviation (in parenthesis) and the mean difference in duration between the vowel pairs. The longest duration is 204.69 msec produced by P9 as she produced /ɑ:/. On contrary, the shortest duration is for /u:/ by P5 with 65.10 msec. In general, P5 produced all the sounds shorter compared to the other participants, ranging from 65.10 msec to 97.20 msec and P1 produced longer durations for most vowels, ranging from 128.10 msec to 171.59 msec. The mean difference between each vowel pair is also different. /ʊ/ and /u:/ have the lowest difference of 3.20 msec, follows by /ɪ/ and /i:/ of 4.33 msec. Surprisingly, /ʌ/ and /ɑ:/ have the highest difference in duration with 49.00 msec. Further analysis is conducted to understand the ability of the participants in producing proper vowel duration.

4.1.2.1 /ɪ/ and /i:/ Pair

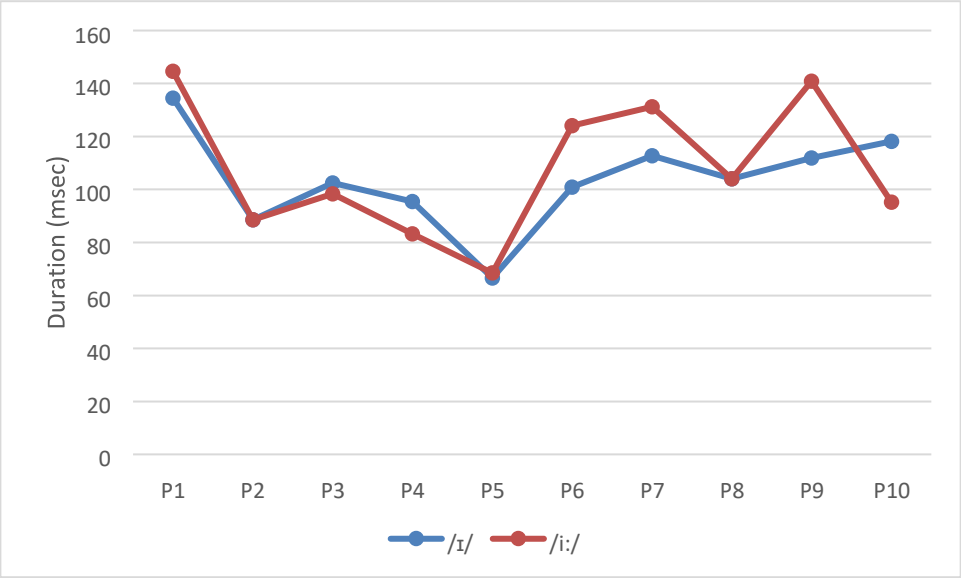

Figure 4.6 Vowel Duration for /ɪ/ and /i:/

Figure 4.6 shows the vowel duration produced by the participants for /ɪ/ and /i:/ pair. The result indicates that the participants were not consistent in producing the length of the sounds, both /ɪ/ and /i:/. The shortest length produced is 55.83 msec for /ɪ/ and 54.67 msec for /i:/. The difference in length between these two sounds is -1.16 msec, indicates that the short vowel /ɪ/ was produced longer than the long vowel /i:/. The longest is 178.97 msec for /ɪ/ and 212.04 msec for /i:/, a difference of 33.07 msec in duration. It is clear that the long /i:/ is produced longer than the short /ɪ/. When the mean of the two data is compared, it shows that some of the participants were not aware of the length difference, as shown in Table 4.3. The mean duration difference explains the difference between the long vowel and short vowel duration produced. The negative values indicate that the participants produced /ɪ/ longer than /i:/. P2, P5 and P8 produced both sounds at the same length, with P5, produced the long vowel slightly longer than the short vowel (1.86 msec) while P8 produced the short vowel slightly longer than the long counterpart (-0.03 msec). P3, P4 and P10 had wrongly produced the long vowel to be shorter than the short vowel (-4.13 msec, -12.29 msec and -22.87 msec respectively). Only P1, P6, P7 and P9 manage to produce the length of the sound properly, but P1 barely did so. Thus, only four out of 10 manage to distinguish the length of long /i:/ and short /ɪ/ properly.

Table 4.3 /ɪ/ and /i:/ Vowel Duration and Mean Different (msec)

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
/ɪ/	134.52	88.59	102.45	95.49	66.62	100.93	112.76	104.01	111.92	118.14
/i:/	144.62	88.59	98.49	83.2	68.48	124.12	131.22	103.98	140.92	95.27
Mean Different	10.1	0	-4.13	-12.29	1.86	23.19	18.46	-0.03	29	- 22.87

4.1.2.2 /ʌ/ and /ɑ:/ Pair

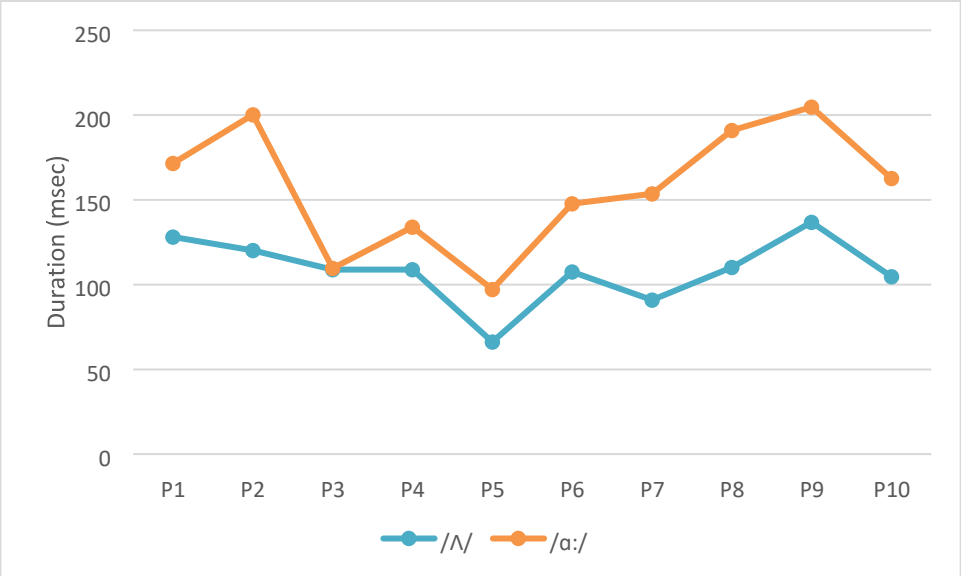

Figure 4.7 Vowel Duration of /ʌ/ and /ɑ:/

Figure 4.7 shows the vowel duration produced by the participants for /ʌ/ and /ɑ:/ pair. The result indicates that the participants were not consistent in producing the length of the sounds, both /ʌ/ and /ɑ:/. The shortest length produced is 52.82 msec for /ʌ/ and 88.15 msec for /ɑ:/. The difference in length between these two sounds is 35.33 msec, which indicates that the short vowel /ʌ/ was produced shorter than the long vowel /ɑ:/. The longest duration is 182.89 msec for /ʌ/ and 290.69 msec for /ɑ:/, a large difference of 107.8 msec in duration. It is clear that the long /ɑ:/ is produced longer than the short /ʌ/. As the two data are compared, the results indicate that the participants are aware of the length difference, as shown in Table 4.4. The mean duration difference explains the difference between the long vowel and short vowel duration produced. All of the participants were able to distinctly produce the length of the vowels accordingly. P3, however, produced a very small difference between short and long vowels. Thus, 10 out of 10 manage to distinguish the length of long and short properly.

Table 4.4 /ʌ/ and /ɑ:/ Vowel Duration and Mean Different (msec)

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
/ʌ/	128.1	120.2	108.88	108.82	66.19	107.55	90.89	110.16	139.83	104.63

/ɑ:/	171.59	200.26	109.67	133.98	97.2	147.8	153.58	190.92	204.69	162.58
Mean Different	43.49	80.06	0.79	25.16	31.01	40.25	62.69	80.86	67.86	57.95

4.1.2.3 /ʊ/ and /u:/ Pair

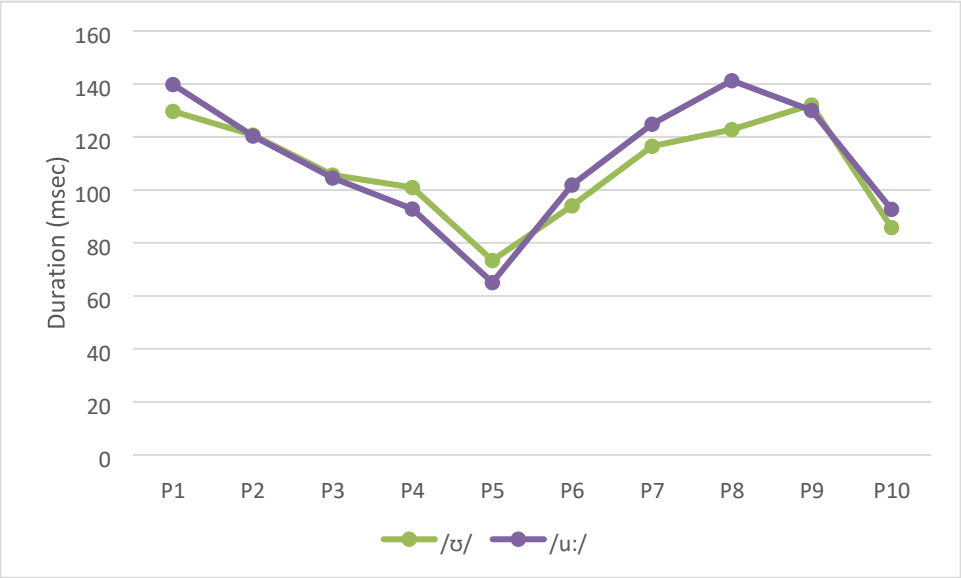

Figure 4.8 Vowel Duration of /ʊ/ and /u:/

Figure 4.8 shows the vowel duration produced by the participants for /ʊ/ and /u:/ pair. The result indicates that the participants were more consistent in producing short vowels but not consistent in producing long vowels. The shortest length produced is 61.79 msec for /ʊ/ and 46.04 msec for /u:/. The difference in length between these two sounds is 15.75 msec, indicates that the short vowel /ʊ/ was produced longer than the long vowel and /u:/. The longest duration is 197.34 msec for /ʊ/ and 229.83 msec for /u:/, a difference of 32.49 msec in duration. It is clear that the long /u:/ is produced longer than the short /ʊ/. As the two data are compared, the results indicate that the participants were not aware of the length difference, as shown in Table 4.5. The mean duration difference explains the difference between the long vowel and short vowel duration produced. The negative values indicate that the participants produced /ʊ/ longer than /u:/. P2, P3 and P9 produced short vowels slightly longer than the long vowel (-0.3 msec, -1.11 msec and -2.05 msec

respectively). In addition, P4 and P5 produced the long /u:/ shorter than the short /ʊ/ (8.12 msec and -8.33 msec). Fortunately, P1, P6, P7, P8 and P10 managed to produce a proper vowel length. Thus, only five out of 10 manage to distinguish the length of long and short properly.

Table 4.5 /ʊ/ and /u:/ Vowel Duration and Mean Different (msec)

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
/ ʊ /	129.64	120.7	105.62	100.92	73.43	94.03	116.5	122.82	132.16	85.78
/ u:/	139.88	120.4	104.51	92.8	65.1	101.87	124.91	141.32	130.11	92.72
Mean Different	10.24	-0.3	-1.11	-8.12	-8.33	7.84	8.41	18.5	-2.05	6.94

4.1.2.4 Analysis of the Findings

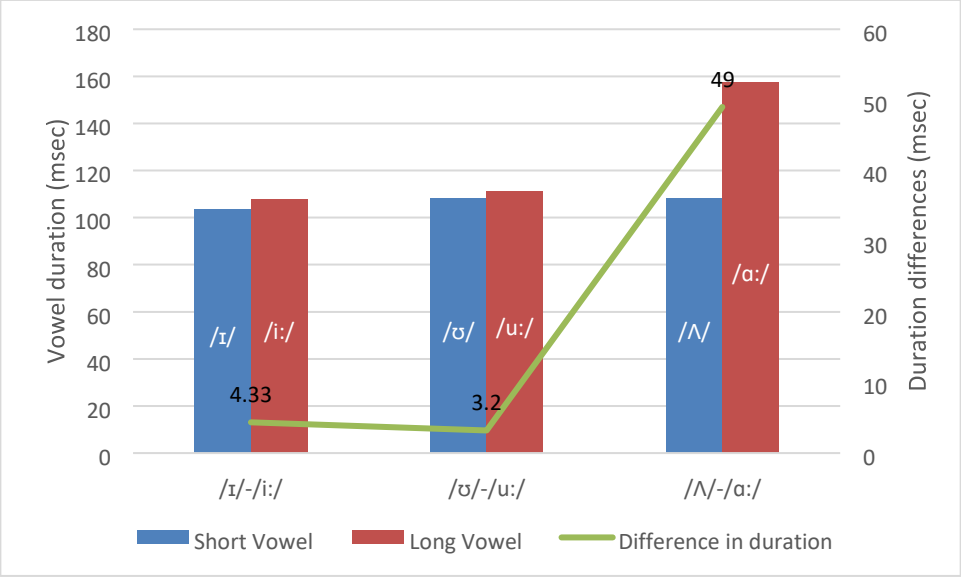

Figure 4.9 Mean Vowel Duration and Difference in Duration (msec)

Figure 4.9 shows the mean vowel duration produced by the participants and the difference in length between the short and long vowel pairs. The /ʊ/ and /u:/ pair has the lowest difference in length between the short and long vowels with 3.2 msec, followed by the /ɪ/ and /i:/ pair, which is at 4.33 msec. /ʌ/ and /ɑ:/ has the highest difference of 49.00 msec, a large gap between the other pair. A paired-sample t-test was conducted to compare the difference in length of all three short and long vowel pairs. Statistical analysis shows

that there is no significant difference in duration of /ɪ and /i:/ (t (59) = 1.01, p=.317). Similar result is produced with /ʊ/ and /u:/ pair as there is no significant difference in duration with t (59) = -.866, p = .390. On the contrary, vowel pair /ʌ/ and /ɑ:/ presents a contrast result as is significant difference in duration of t (59) = -8.440, p = .000. The analysis suggests that the participants are able to discriminate clearly the /ʌ/ and /ɑ:/, as the length for each duration is clearly distinguished.

4.1.3 Summary

To summarise, in terms of vowel quality, the participants produce smaller and compact vowel space areas in contrast with native speakers. The participants are not able to distinguish between /ɪ/ and /i:/ properly, as they swapped the F2 position of these two sounds. Moreover, the distance between the vowel pair is also closer, as an example, the /ʌ/ and /ɑ:/ are produced very closely compared to the other pairs. Thus, further analysis is done in terms of vowel duration produced by the participants, in order to understand their ability to distinguish these vowel pairs.

To summarise, in terms of vowel duration, the participants are aware of the difference in vowel length. Each pair is discriminated, as the short vowels are produced shorter than the longer counterpart. However, the ability to distinguish the duration is different based on the vowel pairs. /ɑ:/ is produced with great length compared to the short vowel /ʌ/, while the /ʊ/ and /u:/ are being discriminated with a small amount of length. Participants were also struggling to properly discriminate /ɪ/ and /i:/, as the difference is also very small. Further analysis of the quality and duration of the targeted vowels is conducted for a better understanding.

4.2 Perception Test

The perception test analysed the ability to distinguish English words containing the targeted sound correctly. The target sounds were embedded into CVC words, and place

in a carrier sentence. A reading of the sentences was recorded. Psychopy software was used to present the reading to the participants, and two cartoon images of the possible answer options for the participants to choose. The result of their responses was analysed and tabulated. The findings are presented by looking at the accuracy and the latency of the responses.

4.2.1 Correct Identification

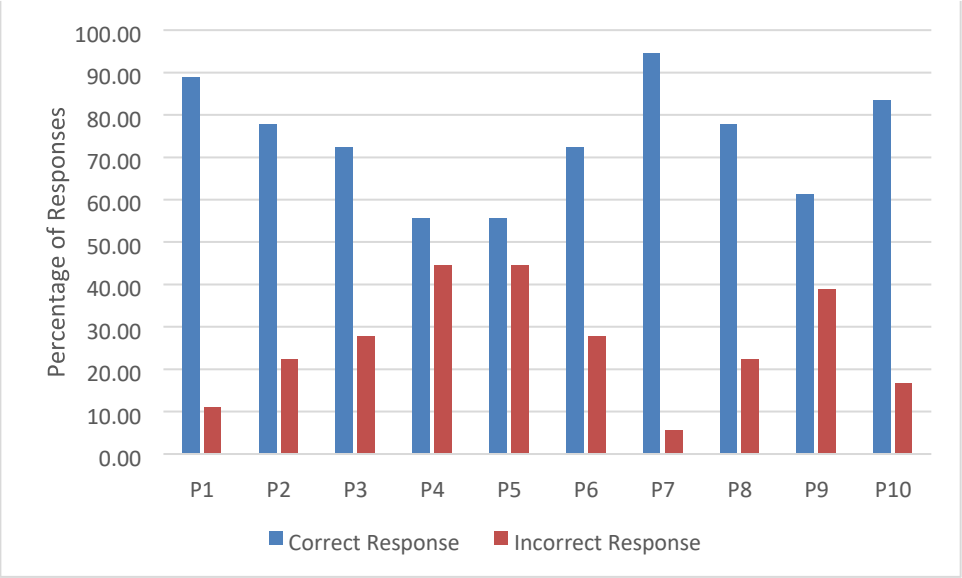

Figure 4.10 Percentage of Participants' Response

Figure 4.10 illustrates the percentage of participant responses, based on the correct and incorrect identification. Participant 7 performed very well with the most correct responses of 94.4%. This participant only chose the incorrect response once. Participants 4 and 5 were tied with 55.6% correct responses and 44.4% incorrect responses, which are the lowest correct responses and highest incorrect responses recorded. In general, all participants were able to achieve at least 50% of correct responses.

4.2.1.1 /ɪ/ and /i:/ Pair

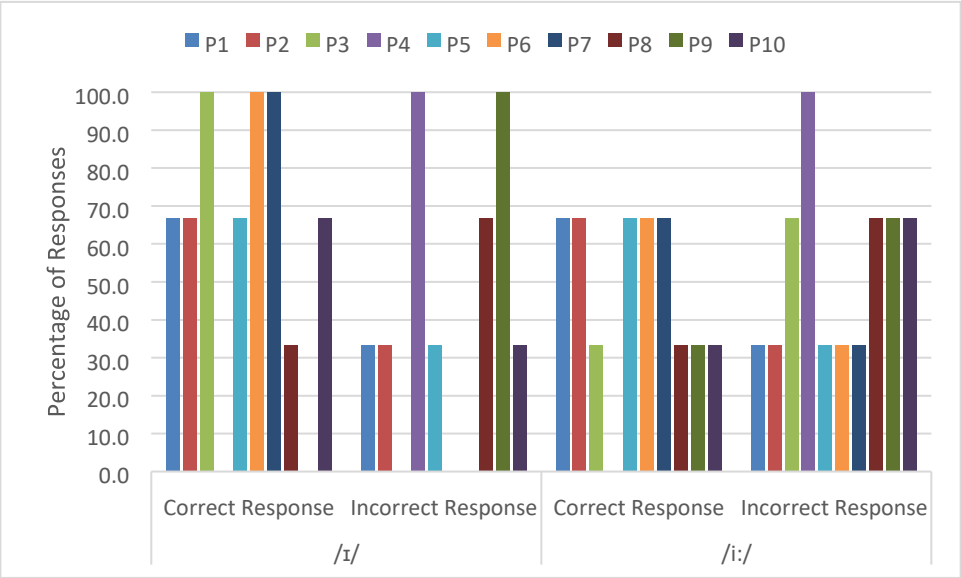

Figure 4.11 Percentage of Responses for /ɪ/ and /i:/ Pair

Figure 4.11 shows the percentage of responses for /ɪ/ and /i:/ pair. In general, the participants were better at perceiving the /ɪ/ than /i:/. P3, P6 and P7 managed to identify /ɪ/ without making an incorrect choice. On contrary, P4 and P9 did not manage to identify /ɪ/ correctly at all. The other participants were able to perceive better by scoring 66.7 % of correct responses, except for P8. As for /i:/, none of the participants scored 100% correct responses. Unfortunately, P4 did not manage to perceive the /i:/ at all, similar to her performance in perceiving /ɪ/ earlier. P1, P2, P5, P6 and P7 managed to score 66.7% of the correct response, in contrast to P3, P8, P9 and P10. For /ɪ/, 7 out of 10 participants can perceive the sound correctly (equal to 70%) while 5 out of 10 participants can perceive /i:/ correctly (equal to 50%). This does support that the participants perform better in perceiving /ɪ/ than /i:/.

4.2.1.2 /Λ/ and /ɑ:/ Pair

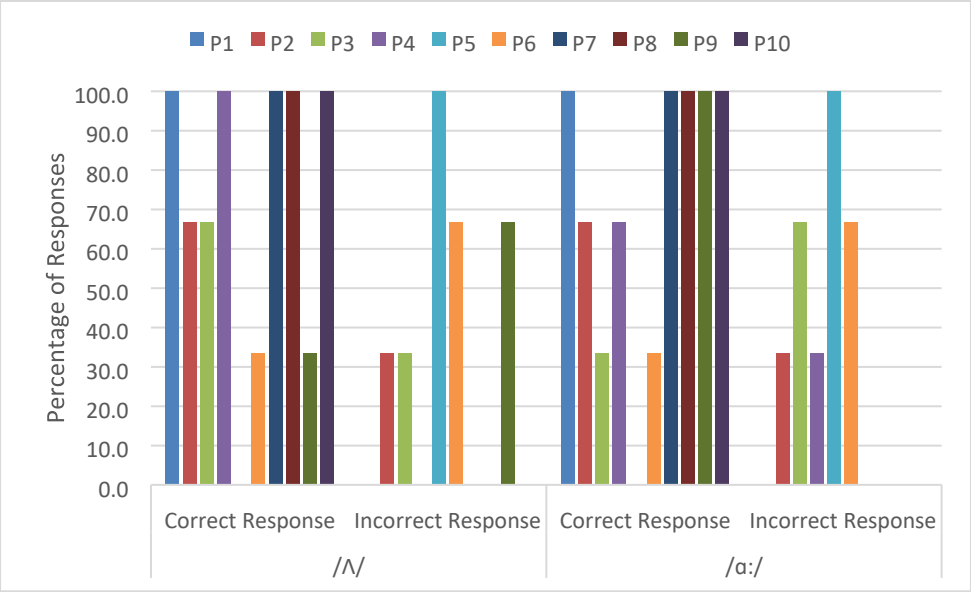

Figure 4.12 Percentage of Responses for /Λ/ and /ɑ:/ pair

Figure 4.12 illustrates the percentage of responses for /Λ/ and /ɑ:/ pair. P1, P4, P7, P8, and P10 were able to perceive /Λ/ with 100% correct responses. However, P5 was not able to do the same as she scored 100% incorrect response. P2 and P3 did better in perceiving this vowel sound as they managed to get 66.7% of the correct response, a contrast to P6 and P6 who scored 33.3% of the correct response. For /ɑ:/, P1, P7, P8, P9 and P10 scored 100% of correct responses. P5 again did not manage to perceive the sound correctly with a 100% incorrect response. P2 and P4 did better than P3 and P6, as the earlier couple scored 66.7% of the correct response and the latter scored 66.7% of the incorrect response. 7 out of 10 participants (70%) performed better in perceiving both /Λ/ and /ɑ:/.

4.2.1.3 /ʊ/ and /u:/ Pair

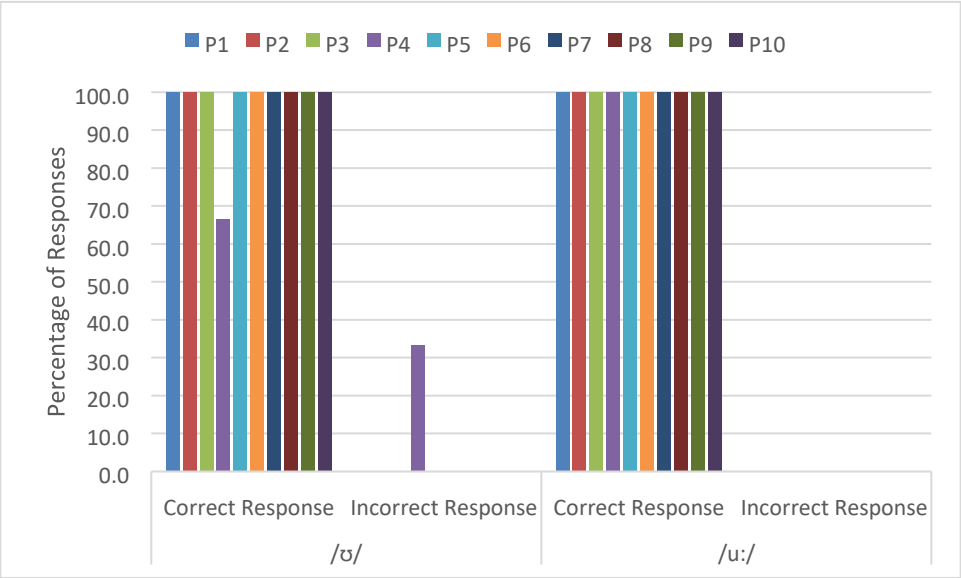

Figure 4.13 Percentage of Responses for /ʊ/ and /u:/ Pair

Figure 4.13 shows the percentage of responses for the /ʊ/ and /u:/ pair. The participants were able to perform better at perceiving /ʊ/ and /u:/. For /ʊ/, apart from P4, all the other participants managed to score 100% correct responses. P4 only scored 66.7% of the correct response. The participants gave even better ability in perceiving /u:/ since all of them scored 100% of correct responses. Thus, 100% of the participants can perceive the vowel sound /ʊ/ and /u:/.

4.2.1.4 Analysis of the Findings

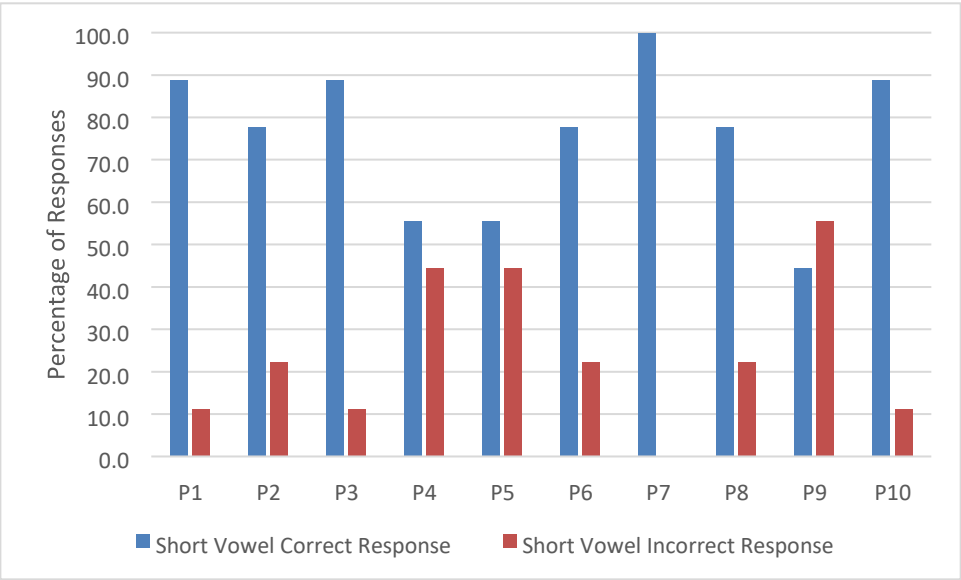

Figure 4.14 Percentage of Responses for Short Vowel

Figure 4.14 shows the percentage of participants' responses to the short vowel. The three short vowels are /ʊ/, /ʌ/ and /ɪ/. P7 did very well in perceiving the short vowel, scoring 100% of the correct response. On contrary, P9 was not performing well with 44.4% of correct responses and 55.6% of incorrect responses. P1, P3 and P10 scored 88.9% of correct responses and 11.1% incorrect response while P4 and P5 perceived the short vowel 55.6% correctly and 44.4% incorrectly. The other two participants, P6 and P8 scored 77.8% correct response and 22.2% of incorrect response. In general, 90% of the participants (9 out of 10) are able to perceive short vowels correctly.

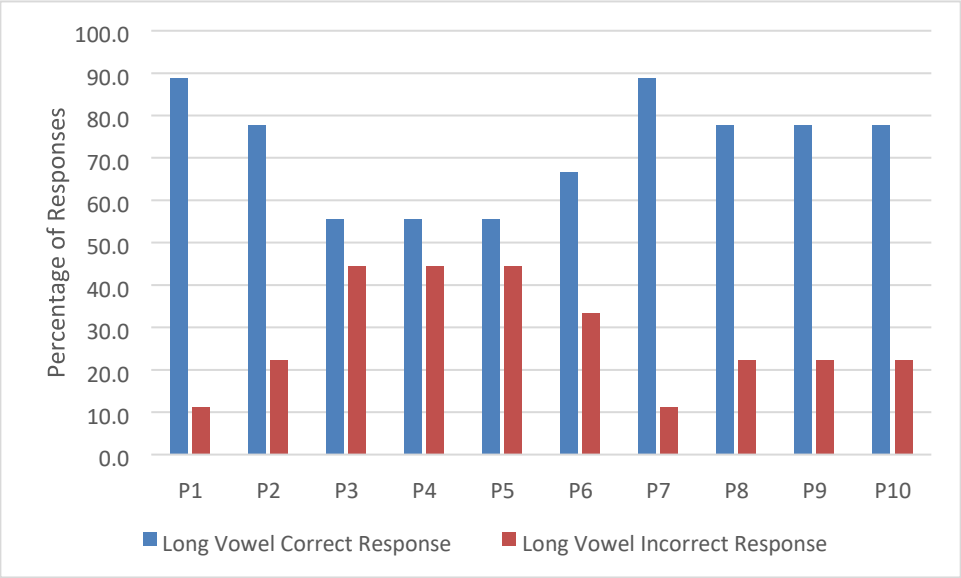

Figure 4.15 Percentage of Responses for Long Vowel

Figure 4.15 shows the percentage of participants' responses for the long vowel, /u:/, /ɑ:/ and /i:/. All participants were able to perceive the long vowel with more correct responses compared to incorrect responses. P1 and P7 scored the highest correct response percentage of 88.9% while the percentage of incorrect response is 11.1%. The lowest score for the correct response was scored by P3, P4 and P5, 55.6% while the incorrect response is at 44.6%. P2, P8, P9 and P10 scored 77.8% of correct response and 22.2% of the incorrect response. Lastly, P6 scored 66.7% on the correct response and 33.3% on the

incorrect response. Overall, 100% of the participants can perceive the long vowel correctly.

Looking at these two figures (Figure 4.17 and Figure 4.18), we can conclude that the participants are better at perceiving the long vowel, compared to the short vowel. Thus, further analysis is conducted by comparing the findings based on the vowel type.

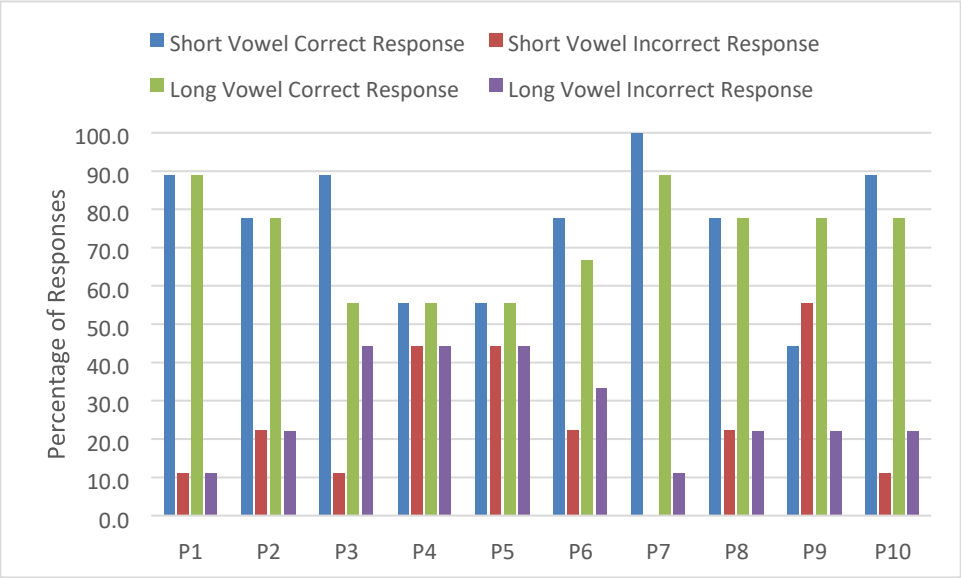

Figure 4.16 Percentage of Responses for Short and Long Vowels by Participants

Figure 4.16 presents the percentage of participants' responses in perceiving short and long vowels. Apart from P9, the other participants scored more than 50% of correct responses, both short and long vowel tasks. P7 did better than the rest with 100% and 88.9% on correctly perceived both vowels. P4 and P5 shared similar results, just like P2 and P8. Thus, the data were then combined and compared based on the short and long vowels.

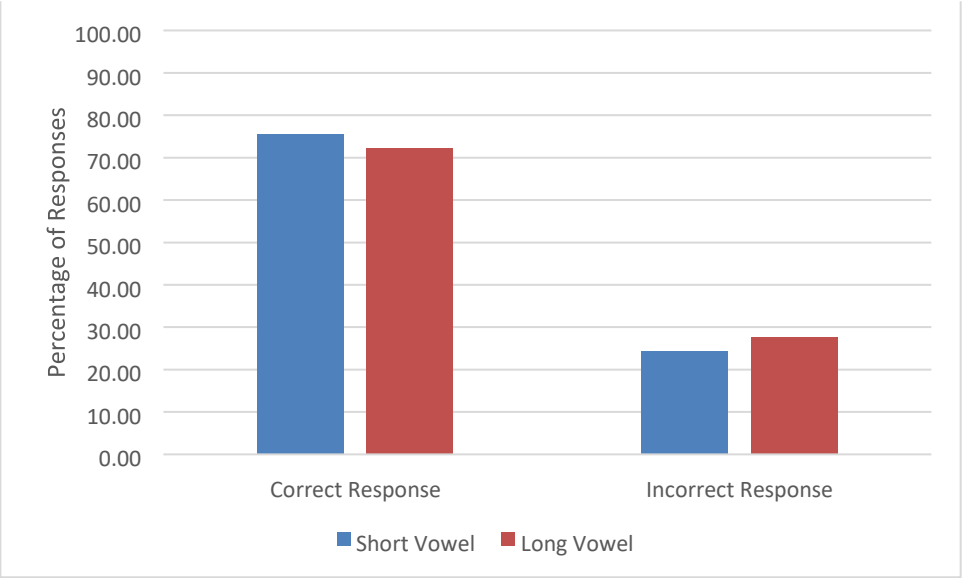

Figure 4.17 Percentage of Responses for Short and Long Vowel

Figure 4.17 shows the percentage of responses for the short and long vowels, for both correct and incorrect responses. Participants managed to score more than 70% correct responses, 75.6% for short vowels and 72.7% for long vowels. Surprisingly, initially, the finding showed that the participants did better in perceiving the long vowel than a short vowel, but further analysis showed otherwise. A paired-sample t-test was conducted to see whether there is a significant difference between the short and long vowel correct responses and no significant difference is found (t (89) = -.555, p = .580).

4.2.2 Reaction Time (Latency)

Table 4.6 Average Reaction Times (in second) of English Monophthong based on Response Accuracy and Standard Deviation (in Parenthesis)

Participant	Reaction Time (s)											
	/ɪ/		/i:/		/ʊ/		/u:/		/ʌ/		/ɑ:/	
	Correct Response	Incorrect Response	Correct Response	Incorrect Response	Correct Response	Incorrect Response	Correct Response	Incorrect Response	Correct Response	Incorrect Response	Correct Response	Incorrect Response

1	1.77 (1.54)	2.93 (5.08)	1.50 (1.30)	1.16 (2.00)	2.57 (0.19)	na	2.37 (0.14)	na	2.55 (0.34)	na	3.11 (0.18)	na
2	1.55 (1.35)	0.85 (1.47)	3.16 (3.40)	2.14 (3.70)	2.58 (0.47)	na	2.70 (0.30)	na	2.08 (1.82)	1.02 (1.76)	1.59 (1.76)	3.93 (6.81)
3	3.11 (0.74)	na	1.35 (2.34)	1.86 (1.61)	2.88 (0.17)	na	3.18 (0.50)	na	1.90 (2.44)	0.84 (1.45)	1.12 (1.93)	2.99 (3.02)
4	na	8.16 (1.80)	na	8.48 (6.80)	2.81 (0.94)	1.18 (2.04)	2.54 (1.04)	na	3.51 (0.39)	na	2.23 (2.91)	1.69 (2.93)
5	5.74 (6.44)	1.35 (2.34)	1.59 (1.40)	1.02 (1.76)	1.61 (1.01)	na	1.46 (0.24)	na	na	2.99 (0.93)	na	2.55 (1.25)
6	7.96 (4.94)	na	1.98 (1.75)	1.14 (1.97)	3.51 (0.04)	na	3.91 (0.80)	na	0.96 (1.66)	1.98 (1.77)	1.14 (1.98)	2.48 (2.17)
7	5.67 (2.64)	na	2.56 (2.23)	1.00 (1.73)	2.00 (1.18)	na	2.29 (1.54)	na	2.67 (0.13)	na	2.45 (0.87)	na
8	1.01 (1.75)	1.74 (1.96)	1.53 (2.64)	2.54 (2.22)	3.50 (0.46)	na	3.52 (0.68)	na	3.46 (0.81)	na	1.17 (2.02)	2.36 (2.06)
9	na	4.78 (1.75)	1.81 (3.14)	3.98 (3.79)	3.62 (0.14)	na	3.62 (0.10)	na	1.74 (3.01)	3.10 (2.76)	4.77 (1.69)	na
10	3.90 (3.42)	1.85 (3.20)	2.07 (3.58)	5.02 (5.39)	3.00 (0.29)	na	3.09 (0.44)	na	3.85 (1.32)	na	4.87 (1.02)	na
Average	3.07 (3.63)	2.17 (3.15)	1.75 (2.16)	2.83 (3.74)	2.81 (0.81)	0.12 (0.64)	2.87 (0.93)	na	2.27 (1.73)	0.99 (1.73)	2.25 (2.08)	1.60 (2.68)

Table 4.6 presents the average reaction time or latency on the English monophthongs by the participants based on the accuracy. The reaction time is measured in second (s). The 'na' indicates that no data is available due to no responses were recorded. The shortest reaction time was recorded at 0.84 s by P3 for /ʌ/ incorrect response. The longest reaction time was by P4 at 8.48 s for /i:/ incorrect response. In general, participants took an average of 0.12 s to 3.07 s to respond to the stimuli, whether by selecting the correct response or the incorrect response. No reaction time was recorded for /u:/ incorrect response due to all participants manage to respond correctly to the stimuli. The next part looks at the average difference between the vowel pair. The average difference of correct response for /ɪ/ and /i:/ pair was at 1.32 s while the incorrect response was at 0.76 s. The average

difference of correct response for Λ/ and /ɑ:/ pair was at 0.03 while the incorrect response was at 0.61 s. The average difference of correct response for the /ʊ/ and /u:/ pair was at 0.06 s while the incorrect response was at 0.12 s. /ɪ/ and /i:/ average reaction time for correct response was the longest compared to the other two pairs while the average reaction time of correct response for Λ/ and /ɑ:/ pair was the shortest. Further analysis was conducted to look into the reaction times of each vowel pair.

4.2.2.1 /ɪ/ and /i:/ Pair

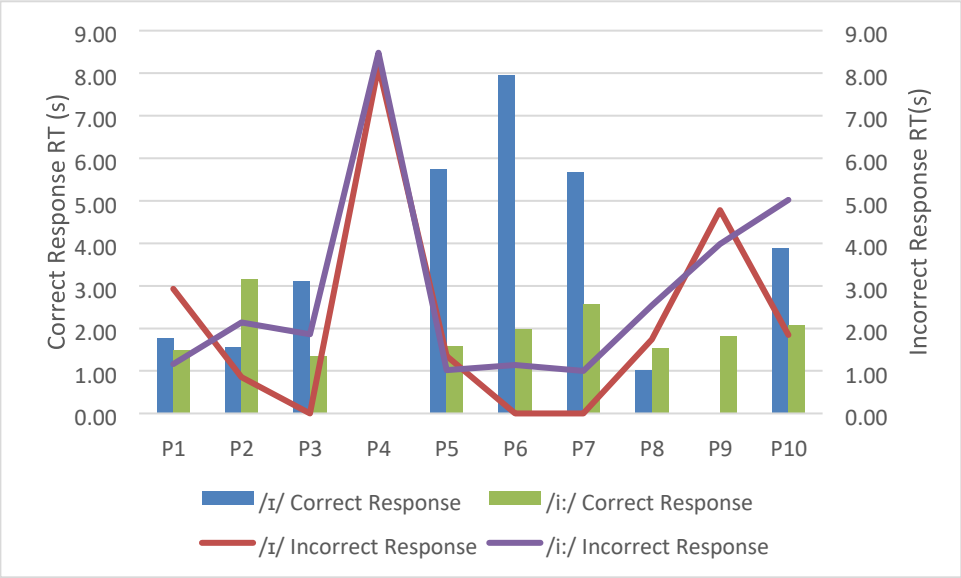

Figure 4.18 Average Reaction Time (s) for /ɪ/ and /i:/ Responses

Figure 4.18 shows the average reaction time for /ɪ/ and /i:/ pair by the participants. P4 took 8.16 s to respond to the stimuli, and her responses were all incorrect. This is the longest time taken to react to /ɪ/. Aside from P4, P5, P6 and P7 also took more than 5 seconds to respond to the stimuli; 5.74 s (correct response), 7.96 s (correct response) and 5.67 s (correct response) respectively. The other participants managed to the reaction below 5 seconds, ranging from 0.85 s (P2 for incorrect response) to 4.78 s (P9 for incorrect response). As for the shortest reaction time, it was by P2 at 0.85 for the incorrect response too. The participants took longer to respond correctly for /ɪ/ at 30.71 s compare to incorrect response at 21.66 s. On the other hand, /i:/ shows an interesting finding. P4 again took the longest to react to the stimuli, at 8.48 s for the incorrect response. The other participant

that took more than 5 seconds to respond to the stimuli was P10, at 5.02 s for the incorrect response. The other participants managed to respond below 5 seconds, ranging from 1.02 (P5 for incorrect response) to 3.98 s (P9 for incorrect response). Overall, the participants took longer to respond to the incorrect response, with a total average of 28.34 s than correct response, at 17.55 s with a difference of 10.79 s.

In total average, the participants took longer to respond correctly to /ɪ/ than /i:/, at 30.71 s and 17.55 s respectively, with a 13.16 s difference. Comparatively, their incorrect responses for /ɪ/ was shorter than /i:/, took 21.66 s and 28.34 s respectively, with a 6.68 s difference. Comparing the reaction time for correct response and incorrect response for both vowels, the participants took a longer time to respond incorrectly, with a total average of 50.0 s while correct response took a total average of 48.26 s. On average, the participants took 2.41 s to respond correctly and 2.50 s to respond incorrectly to /ɪ/ and /i:/ pair.

4.2.2.2 /ʌ/ and /ɑ:/ Pair

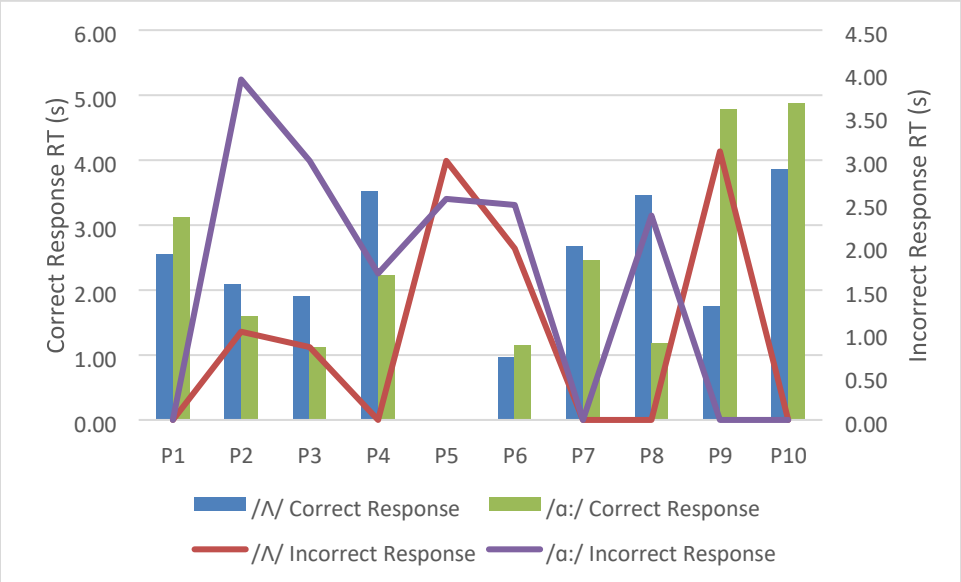

Figure 4.19 Average Reaction Time (s) for /ʌ/ and /ɑ:/ Responses

Figure 4.19 shows the average reaction time for /ʌ/ and /ɑ:/ by the participants. P10 took 3.85 s to respond correctly to /ʌ/, the longest response recorded for both correct and

incorrect response for /ʌ/. The shortest reaction time was taken by 0.84 s by P3 for the incorrect response. No participants took longer than 5 seconds to respond to the stimuli. The participants took longer to respond correctly, with a total average of 22.72 seconds compared to incorrect response at 9.93 s, with a difference of 12.79 s. On the other hand, the longest reaction time recorded for /ɑ:/ was at 4.87 for the correct response by P10, while the shortest reaction time was by P3 at 1.12 s for the correct response. Again, no participants took longer than 5 seconds to respond to the stimuli. In total average, the participants took longer to respond correctly compared to incorrectly, at 22.45 s and 16.0 s respectively, a difference of 6.45 seconds.

In total average, the participants took longer to respond correctly to /ʌ/ and /ɑ:/, at 22.72 s and 22.45 s respectively, with a slightly different of 0.27 s. Comparatively, the reaction time of incorrect responses for /ʌ/ was shorter in total average than /ɑ:/, at 9.93 s and 16.00 s respectively, with a 6.07 s difference. Comparing the reaction time for correct response and incorrect response for both vowels, the participants took a longer time to respond correctly, with a total average of 45.17 s while incorrect response took a total average of 25.93 s. On average, the participants took 3.22 s to respond correctly and 1.30 s to respond incorrectly to /ʌ/ and /ɑ:/ pair.

4.2.2.3 /ʊ/ and /u:/ Pair

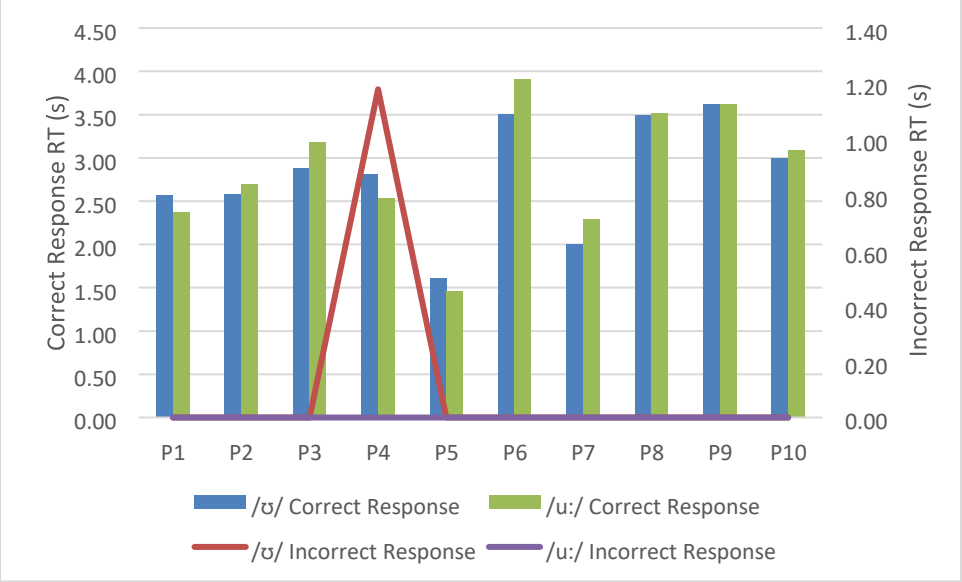

Figure 4.20 Average Reaction Time (s) for /ʊ/ and /u:/ Responses

Figure 4.20 shows the average reaction time for /ʊ/ and /u:/ by the participants. The shortest reaction time recorded was 1.18 s for the incorrect response of /ʊ/ by P4. P9 recorded the longest reaction time for correct response at 3.62 s. The participants took longer to react to the correct response with a total average of 28.08 s as compared to the incorrect response of 1.18 s. No participants took longer than 5 seconds to respond to the stimuli. On the other hand, there were only data of correct response for /u:/ due to no response for incorrect response were collected. P5 took the shortest reaction time of 1.46 s while P6 took the longest reaction time of 3.91 s. No participants took longer than 5 seconds to respond to the stimuli. The participants took 28.68 s in total average to respond correctly.

In total average, the participants took longer to respond correctly to /u:/ than /ʊ/, at 28.68 s and 28.08 s respectively, with a slight difference of 0.60 s. Comparatively, the reaction time of incorrect responses for /ʊ/ was 1.18 s and no data for /u:/. Obviously, the reaction time for correct response was longer than incorrect response, with a total average of 56.76 s while incorrect response took a total average of 1.18 s. On average, the

participants took 2.84 s to respond correctly and 0.06 s to respond incorrectly to /u:/ and /ʊ/ pair.

4.2.2.4 Analysis of the Findings

Most of the participants did respond within 5 seconds after the stimuli were presented to them, whether the responses were correct or not, except for P4, P5, P6 and P7. These participants were struggling to respond quickly to /ɪ/ and /i:/, but not towards the other vowels. P1 and P3 demonstrated a good time reflect by responding to the stimulus in less than 3.2 seconds. In a total average of the reaction time, P4 took the longest to respond to the stimuli at 30.60 s while P1 produced the shortest reaction time at 17.96 s. On average, the participants' responses took 1.50 s to 2.55 s to be recorded.

Figure 4.21 Total and Average Reaction Time (s) based on Vowel.

Figure 4.21 shows the total and average reaction time based on vowels. Participants took the longest to respond correctly to /ɪ/ as they took a total of 92.1 s, at an average of 3.07 s. These were followed by the correct response /u:/ and incorrect response /i:/, 86.04 s (at an average of 2.87 s) and 85.0 s (at an average of 2.83 s) respectively. The total reaction time of correct response for /ʊ/ was at 81.42 s, /ʌ/ at 68.18 s and /ɑ:/ at 67.35 s,

and their average reaction time were at 2.81 s, 2.27 s and 2.25 s respectively. In addition, the total reaction time of incorrect response for /ɪ/ was at 64.96 s (average at 2.17 s), correct response for /i:/ at 52.63 s (average at 1.75 s), incorrect response for /ɑ:/ was at 48.01 s (average at 1.60 s) and incorrect response for /ʌ/ was at 29.78 s (average at 0.99 s). The incorrect response for /ʊ/ took a total of 3.53 s and at an average of 0.12 s, but this is due to only one response was recorded. No data was recorded for incorrect response /u:/, thus the total reaction time and average reaction were marked at zero. Table 4.7 shows the total average reaction time, difference in average reaction time between correct response and incorrect response, average reaction time and difference in average reaction time between correct response and incorrect response (s). The participants took more time to react to the correct response, except for vowel /i:/. The difference in total reaction time shows a negative value as the participants took 32.37 s longer to react to incorrect response, at different in average reaction time of 1.08 s. A paired-sample t-test was conducted to see the significant difference between the reaction time of the correct response and the incorrect response of the vowel. There are no significant between correct and incorrect response for /ɪ/ (t (29) = .817, p = .421), /i:/ ((t (29) = -1.099, p = .281) and /ɑ:/ (t (29) = .811, p = .424). However, the is a significant different in /ʌ/ (t (29) = 2.184, p = .037). Due to lack of data, the /u:/ and /ʊ/ pair was not statistically analysed. The data was then further analysed by comparing each of the vowel pairs.

Table 4.7 Total Average Reaction Time, Different in Total Average Reaction Time, Average Reaction Time and Different in Average Reaction Time of the Targeted Vowels (s)

Vowel	Responses	Total Average Reaction Time	Different in Total Reaction Time	Average Reaction Time	Different in Average Reaction Time
/ɪ/	Correct Response	92.10	27.14	3.07	0.90
	Incorrect Response	64.96		2.17	
/i:/	Correct Response	52.63	-32.37	1.75	-1.08

	Incorrect Response	85.00		2.83	
/ʊ/	Correct Response	81.42	77.89	2.81	2.69
	Incorrect Response	3.53		0.12	
/u:/	Correct Response	86.04	na	2.87	na
	Incorrect Response	0.00		0.00	
/ʌ/	Correct Response	68.18	38.40	2.27	1.28
	Incorrect Response	29.78		0.99	
/ɑ:/	Correct Response	67.35	19.34	2.25	0.64
	Incorrect Response	48.01		1.60	

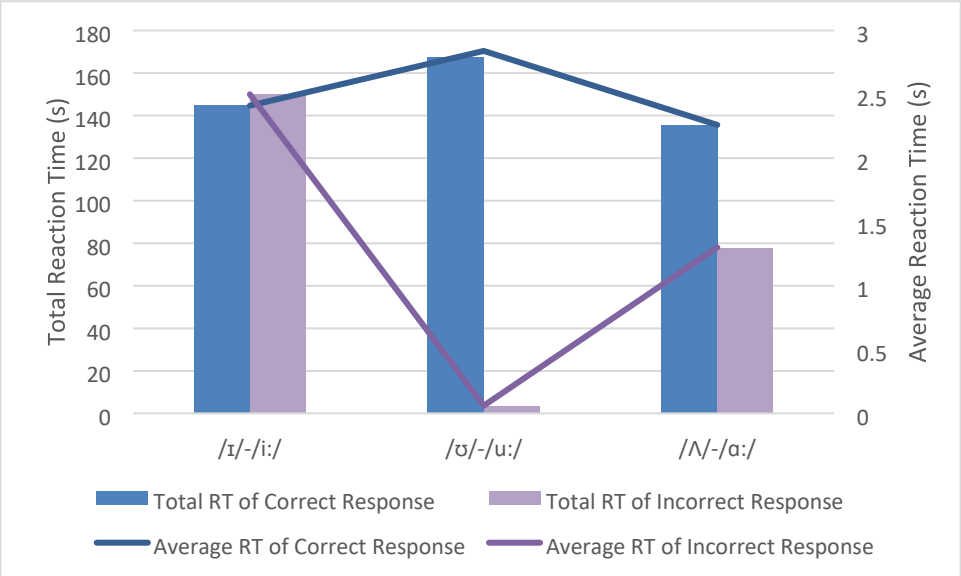

Figure 4.22 Total and Average Reaction Time based on Vowel Pairs

Figure 4.22 presents the total reaction time and average reaction time on the stimuli based on vowel pairs. The total reaction time of the incorrect response for /ɪ/ and /i:/ was at 149.96 s, slightly longer than the correct response at 144.73 s, a difference of 5.23 s. As for the average reaction time, the correct response for /ɪ/ and /i:/ was at 2.41 s, compared to the incorrect response at 2.50 s. The /ʌ/ and /ɑ:/ pair has a wider gap between the reaction time of correct response and incorrect response. The total reaction time taken for the correct response was at 135.53 s and the incorrect response was at 77.79 s, with a difference of 57.74 s. Comparing to the average reaction time for this pair, the correct response was recorded at 2.26 s and the incorrect response was recorded at 1.30 s. Lastly,

the /ʊ/ and /u:/ pair. The total reaction time for correct response was at 167.46 s, in comparison to the incorrect response at 3.53 s, a huge difference at 163.93 s. As for the average reaction time, the correct response recorded 2.84 s and the incorrect response recorded 0.06 s. Putting the values of the average reaction time of the participants towards the stimuli in order, the correct response for /ʊ/ and /u:/ pair took the longest time, followed by the incorrect response for /ɪ/ and /i:/, correct response for /ɪ/ and /i:/, and correct response for /ʌ/ and /ɑ:/. Next was the incorrect response for /ʌ/ and /ɑ:/ and lastly, was the incorrect response for /ʊ/ and /u:/.

A paired-sample t-test is conducted to determine the significant different of total reaction time of correct responses and incorrect responses between each pair. Vowels /ɪ/ and /i:/ produce no significant different in correct response ($t(29) = 2.023$, $p = .052$). Similarly, vowels /ʌ/ and /ɑ:/ produce no significant different in correct response with $t(29) = .067$, $p = .947$, as well as the pair /ʊ/ and /u:/; also show no different result in the correct response ($t(29) = -.980$, $p = .335$). Similar results were also produced for the incorrect responses of the pairs. Vowels /ɪ/ and /i:/ produce no significant different in incorrect response ($t(29) = -1.035$, $p = .309$). Similarly, vowels /ʌ/ and /ɑ:/ produce no significant different in incorrect response with $t(29) = 1.000$, $p = .326$. Lastly, the pair /ʊ/ and /u:/ also show no different result in the incorrect response ($t(29) = -1.257$, $p = .218$).

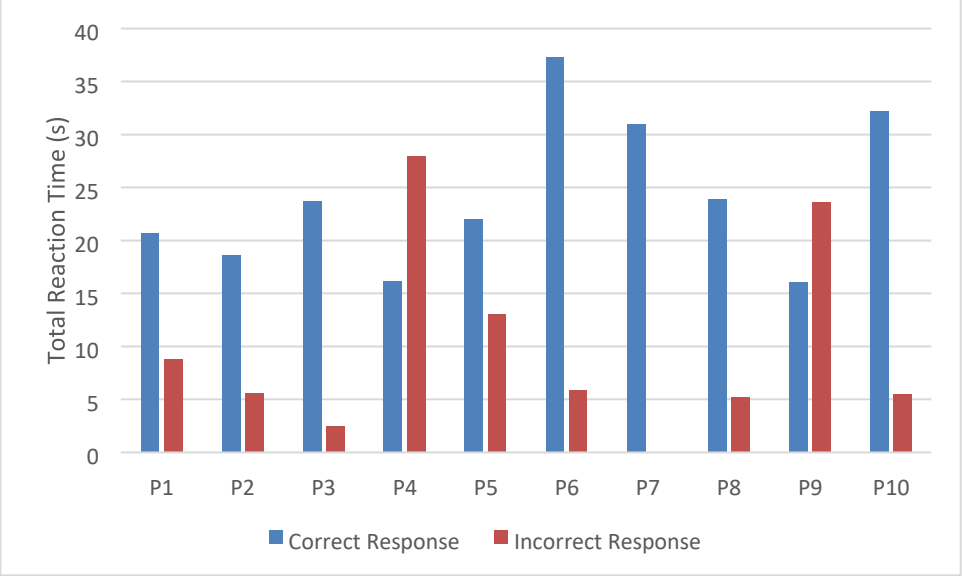

Figure 4.23 Short Vowels Total Reaction Time (s) by Participants

Figure 4.23 shows the total time taken by the participants to react to short vowels, both correctly and incorrectly. P6 took the longest time to react correctly to the stimuli given, with a total of 37.29 s. This is followed by P10 (32.27 s), P7 (31 s), P8 (23.89 s), P3, 23.67 s), P5 (22.05 s) and P1 (20.67 s). The rest of the participants took less than 20 s to respond correctly: P2 (18.65 s), P4 (16.14 s) and P9 16.07 s). As for the incorrect response, P4 took the longest by taking 28 s. This is followed by P9 (23.63 s) and P5 (13.04 s). The rest of the participants took less than 10 seconds to respond incorrectly: P1 (8.8 s), P2 (5.59 s), P3 (2.52 s), P6 (5.93 s), P8 (5.22 s) and P10 (5.53 s). No data was recorded for P7 as the participant did not produce any incorrect response. Table 4.8 also includes the difference in reaction times (s) by the participants between the correct response and incorrect response. P4 and P9 took more time to produce incorrect responses than correct responses, 11.86 s and 7.56 s respectively. In contrast, the rest of the participants spent more time on correct responses. P6 produced the highest difference between correct response and incorrect response for short vowels. 31.36 s. This is followed by P7 with 31 s, P10 (26.73 s) and P3 (21.15 s). Except for P5 who took 9.01 s, the rest of the participants produced different reaction times between the correct and incorrect responses below 20

seconds but more than 10 seconds. P1 produced a difference of 11.87 s, P2 (12.06 s) and P8 (18,67 s).

Table 4.8 Participants' Total Reaction Time (s) and Different in Reaction Time for Short Vowels

Participant	Correct Response	Incorrect Response	Different in Reaction Time
P1	20.67	8.8	11.87
P2	18.65	5.59	13.06
P3	23.67	2.52	21.15
P4	16.14	28	-11.86
P5	22.05	13.04	9.01
P6	37.29	5.93	31.36
P7	31	0	31
P8	23.89	5.22	18.67
P9	16.07	23.63	-7.56
P10	32.27	5.54	26.73

Figure 4.24 shows the total reaction time taken by the participants to respond to long vowels. P9 took the longest time to react correctly to the stimuli given, with a total of 30.6 s. This is followed by P10 at 30.08 s. P1, P2, P6 and P7 took more than 20 seconds to respond correctly: 20.93 s, 22.34 s, 21.12 s and 21.89 s respectively. Only P5 took less than 10 seconds to respond correctly to the long vowel stimuli, of 9.16 s. The rest of the participants took more than 10 seconds to respond correctly but less than 20 seconds. P3 took 16.96 s, P4 (14.32 s) and P8 (18.63 s).

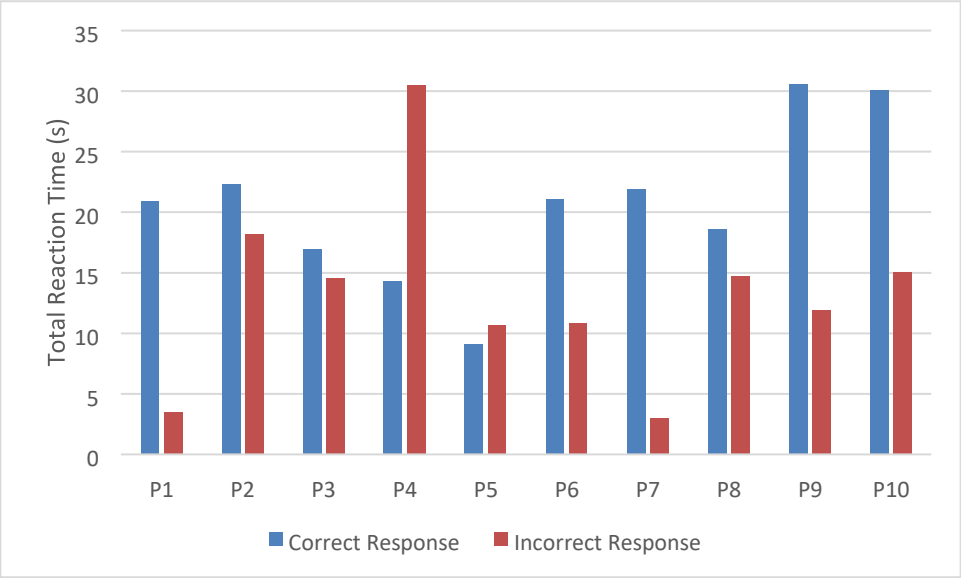

Figure 4.24 Long Vowels Total Reaction Time (s) by Participants

Table 4.9 includes the difference in reaction time between correct responses and incorrect responses. P4 and P5 spent more time on the incorrect response, thus the negative values in the table. Both participants produced a difference of 16.2 s and 1.54 s, respectively. Other participants produced a difference of fewer than 20 seconds between correct and incorrect responses. The highest difference was produced by P7 at 18.89 s, followed by P9 at 18.67 s. P1, P6 and P10's differences in reaction time are at 17.46 s, 10.26s and 15.01 s respectively. Other than that, the different produced were below 10 seconds: P2 (4.13 s), P3 (2.4 s) and P8 (3.92 s).

Table 4.9 Participants' Total Reaction Time (s) and Different in Reaction Time for Long Vowels

Participant	Correct Response	Incorrect Response	Different in Reaction Time
P1	20.93	3.47	17.46
P2	22.34	18.21	4.13
P3	16.95	14.55	2.4
P4	14.32	30.52	-16.2
P5	9.16	10.7	-1.54

P6	21.12	10.86	10.26
P7	21.89	3	18.89
P8	18.63	14.7	3.93
P9	30.6	11.93	18.67
P10	30.08	15.07	15.01

Figure 4.25 presents the total and average reaction times of correct response and incorrect response based on the vowel types, short and long vowels. The participants took a total of 241.7 s to respond correctly to the short vowel and 206.02 s for the long vowel. On average, they took 2.72 s to respond correctly to the short vowel, in comparison with the long vowel (2.29 s). The difference produced is 0.43 seconds. On the other hand, the participants took a total of 98.27 s for the incorrect response of the short vowel, and 133.01 s for the long vowel. As for the average reaction time, the short vowel is at 1.09 s and the long vowel is at 1.48 s, a difference of 0.39 seconds.

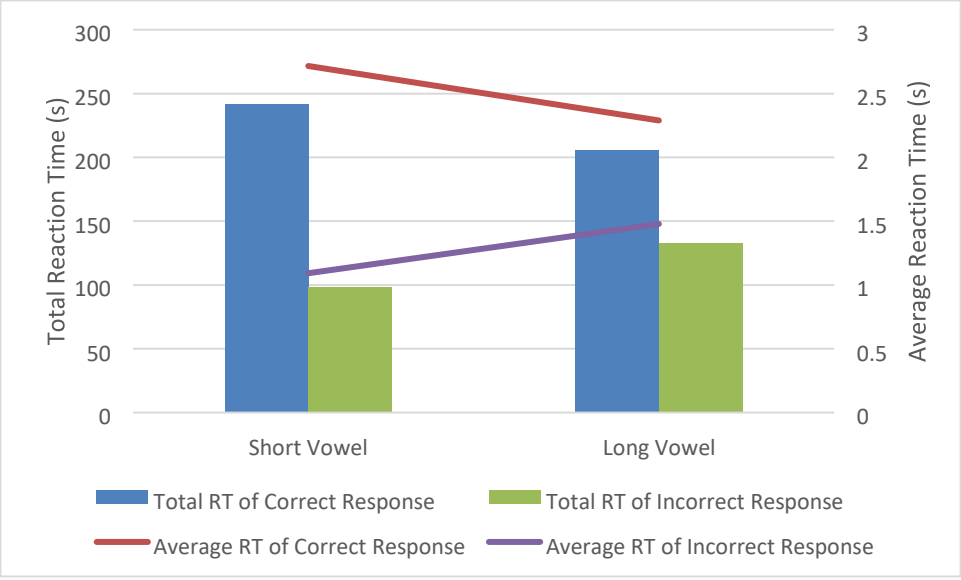

Figure 4.25 Total and Average Reaction Times of Correct and Incorrect Responses for Short and Long Vowels (s)

A paired-sample t-test was conducted to identify the significant differences between the average reaction times of correct responses and incorrect responses of both short and long vowels. For short vowel, there is a significant different in average reaction time for correct response and incorrect responses ($t(89) = 3.718$, $p = .000$). However, a different

result is found for the long vowel, where there is no significant difference in average reaction time between the correct responses and incorrect responses (t (89) = 1.788, p = .077). Further analysis is conducted to investigate the significant difference in average reaction time of the vowel type based on the identification. There is no significant difference between the correct responses of short and long vowels (t (89) = 1.477, p = .143). A similar result is found for the incorrect responses as there is no significant difference between the short and long vowel (t (89) = -1.424, p = .158)

4.2.3 Summary

In terms of vowel accuracy, the participants can perceive and distinguish the vowel accurately, in which they can perceive short vowels better than a long vowel, but with not much significant difference. Looking at vowel pair, /ʊ/ and /u:/ pair is easier to perceive, follow by /ʌ/ and /ɑ:/ pair and lastly the /ɪ/ than /i:/ pair. The /ɪ/ than /i:/ pair is the hardest to perceive as the finding showed that participants were struggling to distinguish the two pairs. Putting the vowel in a hierarchy, the participants can perceive /u:/ the best, follows by /ʊ/, either /ʌ/ and /ɑ:/ in the third place as they shared the same outcome, then /ɪ/ and lastly /i:/.

In terms of vowel latency, the participants spent less than 5 seconds responding to the stimuli presented to them. They allocated more time to correct responses than incorrect responses, but this may be due to the difference in the percentage of the responses. Comparing the vowel types, the participants can react faster to short vowels compared to long vowels, as the paired-sample t-test identified a significant difference for correct responses. However, when comparing short and long vowels to the correct identification, there is no significant difference is found, both correct and incorrect responses.

4.3 Discussion

4.3.1 Production Test

If this result of this research (Figure 4.26) is to compare with the English vowel chart in Chapter 2 (Figure 2.1), the position of each vowel is not accurate. The participants' result shows the position of /ɪ/ is the most left, but lower than /iː/. The position of these two vowels is wrong on the F2 level. This occurrence is also similar to the /ʊ/ - /uː/ pair, in which the position of /ʊ/ is to the right-side of /uː/, also swapped their places. As for the F1 level, these two vowels are positioned in an almost similar plane. Dissimilarly, the F2 of /ʌ/ is at the centre of the chart and /ɑː/ is positioned to its right. However, the F1 level of these two sounds is almost identical. Nonetheless, no speaker, even a native speaker is able to speak exactly as how the standard dictates. In order to analyse the data properly, the participants' vowel plotting is compared to Deterding's vowel plotting.

Figure 4.6 shows the comparison of vowel plotting between research's young female participants versus female participants in Deterding (1997). His participants did not accurately produce the vowels according to the standard vowel chart, but more as normal native speakers speak the language. Participants from both researches have a similar tendency in producing English vowels. Aside from /ɪ/ - /iː/ pair, both /ʊ/ - /uː/ and /ʌ/ - /ɑː/ pairs' F2 positions are similar, /uː/ and /ʌ/ are more front than their counterpart. Moreover, /ʌ/ F1 value is higher than /ɑː/, causing its position to be lower, in contrast to the position of /ʌ/ in the standard vowel chart. Generally, it is typical for speakers to produce this vowel as such.

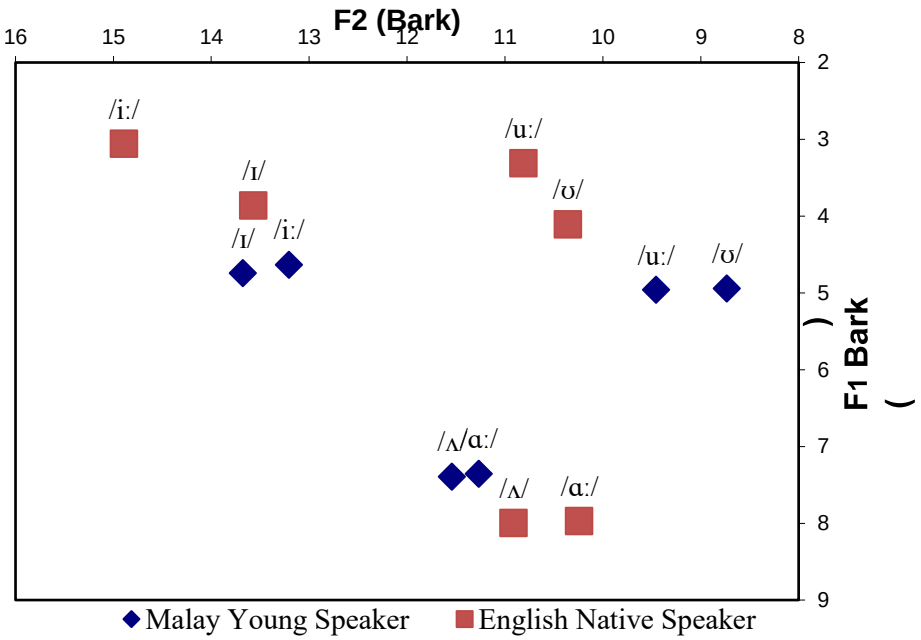

Figure 4.26 Vowel Plotting Malay Young Speakers vs Native Speakers (Deterding, 1997)

Here, it is clearer that the participants' vowel space is smaller compared to the native speakers. In Deterding, the ED of /ʊ/ is the shortest from the centroid of 1.92 Bark and /i:/ is the longest from the centroid of 4.26 Bark. The vowels produced were more towards the centre of the mouth. /ɪ/ - /i:/ and /ʊ/ - /u:/ pairs were produced more backward and downward, in comparison with native speakers. /ʌ/ - /ɑ:/ pair was produced more upwards and inward, in contrast to the native speakers. The result of the above analysis shows that the young Malay speakers produced English monophthong differently than native speakers. Their vowel space is more compact and more back. This is consistent with previous research conducted by (Pillai, 2014; Pillai et al., 2010; Thai et al., 2010). Other Austronesian language speakers also producing a similar result, such as Javanese and Sundanese speakers (Perwitasari et al., 2016).

SLM (Flege, n.d., 1995a) claimed that L2 sound that is too similar to L1 is difficult to acquire, due to the inability to create a new category. Many L2 speakers have the tendency to conflate the vowel pair, as found in Pillai & Delavari (2012). The pair /ɪ/ - /i:/ is the most difficult to produce. In this experiment, the F1 values of both vowels were almost

similar and the position of /i:/ is way back. This is different from the finding in Casillas (2015), in which he found that early learners' production of /i/ does not differ from native speakers. However, it is important to highly that the participants in his study were living in an English-speaking country, in contrast with this research's participants. /ʌ/ - /ɑ:/ pair also produced a similar result. Only /ʊ/ and /u:/ pair produced different F2 values, but not so for their F1 values. To further understand this situation, it is best to look at the position of Malay vowels in the vowel quadrilateral. The closest example is by Clynes & Deterding (2011) work, they mapped Standard Malay Brunei following the standard vowel chart (Figure 2.2). Malay /i/ and /u/ are both located at the upper part of the diagram, similar to English vowels. The F1 position of these two vowels are almost similar, the upper vowels are produced almost at the same F1 value. When the long pair of English vowels are introduced, Malay speakers are not aware of these differences. Thus, this explains why the participants were unable to significantly produce different F1 values.

Furthermore, let's look at the word list. The words were chosen in this research to represent the sound also contributed to the challenge in distinguishing the sounds apart. The words were chosen to represent /ɪ/ and /i:/ are pick and peak, which a minimal pair. Minimal pair is defined as two linguistic units that differ in a single distinctive feature or constituent (Merriam-Webster, n.d.), while bid and beep are not. /u:/ and /ʊ/ were represented by book – boot and cook – coop, and these word pairs are not minimal pairs. The chosen words to represent /ʌ/ and /ɑ:/ are cut – cart and duck and dark. The words representing /ɪ/ - /i:/ and /ʌ/ - /ɑ:/ pairs are minimal pairs. Since the words are almost identical, the participants are not aware of the vowel quality differences. Not surprisingly, they were able to produce different values of F2 for /ʊ/ and /u:/, since the words standing for the pair were not minimal pair; book-boot. As discussed in Chapter 3, some of the words chosen are not minimal pairs were due to the participants' proficiency. The words

chosen are words that have been introduced to them in class as well as based on the academic syllabus.

Contrary to previous research such as Pillai et al. (2010) and Salwani (2005), the participants of this research managed to contrast the duration between the vowel pairs, although not distinctly. /ɪ/ and /i:/ again are the hardest to produce in terms of vowel duration with a very small difference. The words were chosen in this research to represent the sound also contributed to the challenge in distinguishing the sounds apart. The word pick and peak are minimal pairs while bid and beep are not. This may influence the participants' performance. Contrary, the participants performance on /u:/ and /ʊ/ pair are better than the /ɪ/ and /i:/ pair. Differently from the previous words, the word pairs for [u] are not minimal pairs. In /ʌ/ and /ɑ:/ case, both word pairs are minimal pairs. The findings showed that the /ɑ:/ and /u:/ were produced longer compared to their short counterparts/ʌ/ and /ʊ/, respectively. Surprisingly, the production of /ɑ:/ was distinctively longer than /ʌ/, the longest compared to the other long vowels in this experiment. Understandably, Malay does not have long vowels; thus, the idea of vowel duration does not exist. This knowledge illustrates the difficulties to produce long vowels among Malay speakers. However, the /ʌ/ - /ɑ:/ pair is a special case. The reason for the great performance of the participants was due to the orthographic cue, the spelling of the targeted word. The long vowel was represented by the words cart and dark. /r/ is an approximant, consonant that has very little obstruction to the airflow. The characteristic is almost vowel-like, more common in American English because the American accent emphasized the /r/, contrary to British English practice. To new learners, Malay speakers, for example, tend to pronounce every letter as to how they do when speaking in their mother tongue. The sound may interfere with the production of the /ɑ:/, subconsciously increasing the length of the vowel. Other than that, as part of the syllabus, some teachers do explicitly teach their students to lengthen the 'a' sound, if the word is spelt with 'ar'. Hence, this justifies the

longer duration of the /ɑ:/. The implementation of the new syllabus in KSSR may also contribute to the improvement in the production of vowel sounds. The learning standard for the Listening and Speaking Module emphasized the ability to speak with correct pronunciation, stress, rhythm and intonation; throughout the six years of the primary school period. Thus, the constant exposure may assist the students to unconsciously familiarise themselves with the English sound system.

Table 4.10 illustrates the result of the analysis. After analysing the result, it is fair to rank the vowel pair production based on the difficulty to produce. The participants were able to appropriately produce difference /ʊ/ and /u:/ sounds, making them the easiest pair to produced. The pair is slightly discriminate on F1 but clearly discriminated on F2. Although the duration difference is not too wide, the distinction is present. /ʌ/ - /ɑ:/ pair is produced clearly in terms of duration, but less well in terms of vowel quality. Lastly, /ɪ/ - /i:/ pair is the hardest to produce, both vowel quality and duration. It is wise to rank the /ɪ/ - /i:/ pair as the most difficult to produce, follow by /ʊ/ - /u:/ and /ʌ/ - /ɑ:/ pair. Thus, let's look at the result of the perception experiment first, in order to understand the result of the production experiment.

Table 4.10 Average value of F1 & F2 (Bark), average duration (msec) and rank on production difficulties of the targeted vowels

Vowel	/ɪ/	/i:/	/ʊ/	/u:/	/ʌ/	/ɑ:/
F1 (Bark)	4.733	4.621	4.934	4.947	7.373	7.3392
	Slight Different		Slightest Different		Slight Different	
F2 (Bark)	13.278	12.757	8.686	9.376	11.435	11.238
	Slight Different		Slight Different		Slightest Different	
Duration	103.54	107.87	108.16	111.36	108.23	157.23
	4.33 (Medium difference)		3.20 (Minor difference)		49.00 (major difference)	

*Rank on Production Level	1	2	3

**The pairs are ranked from the most difficult to clearly produced (1) to the easier to clearly produce (3).*

4.3.2 Perception Test

Casillas (2015), Kondaurova & Francis (2008) and Perwitasari et al. (2015) suggested that L2 speakers have a tendency to rely on duration cues when listening to English words. This will help them in understanding the message and choosing the select response. Although this duration discrimination may not be utilised in their own speaking (production of English sounds), that does not mean the speakers have no knowledge at all. In addition, Ammar et al. (2016) found that the speakers' language proficiency affects their perception too, as to how the Iraqi speakers in their research. The perception is measured in two aspects, accuracy as to how well they correctly identify the vowel sounds and latency as how fast they think and choose correct answers.

In terms of accuracy, the participants were not able to identify /ɪ/ - /i:/ clearly. Other research such as Ammar et al. (2016), Escudero & Boersma (2004) and Yang & Fox (2014) found that second-language speakers are having difficulties in perceiving the long and short vowels. Even if the speakers are bilingual, they too face difficulties in perceiving the monophthongs correctly (Thai et al., 2010). This result agrees with PAM

(Best & Tyler, 2007) Single Category as the speaker assimilated the L2 sounds into one L1 category and SLM (Flege, 1995a) on the difficulties to acquire L2 sound that is too similar to L1 as this prevents them from creating a new category. However, the participants performed very well in distinguishing /ʊ/ - /u:/, followed by /ʌ/ - /ɑ:/. Both /ʊ/ - /u:/ and /ʌ/ - /ɑ:/ pairs fall under PAM Category Goodness, where the L2 sounds are assimilated into the same L1 category, but one is slightly different from the other. Thus,

increases the ability to distinguish them. Vowel quality-wise, the vowel /ʌ/ and /ɑ:/ are not similar to their equivalent in the Malay language. Malay /a/ is produced at lower front, different than /ʌ/ (mid-low) and /ɑ:/ (lower back).

In terms of latency, the participants took a shorter time to respond correctly compared to the incorrect response. The shortest reaction time is 0.96s (correct response of /ʌ/) and the longest average reaction time is 8.84s (wrong response of /i:/). On average, the participants took 3.42s and 4.86s to respond correctly and incorrectly, respectively. Response latency explains that more accurate responses take a shorter time (Schiepers, 1980). It is a part of human nature to select positive and right answers immediately, as mention by Dunning & Perretta (2002). Table 4.11 shows the accuracy and latency of each vowel and the reaction time to respond. The participants spend less time responding correctly to the vowel than responding incorrectly. They took about 5.12 seconds to respond correctly to /ɪ/, 0.29 seconds faster than incorrect response (5.41 s). The reaction time for correct response (3.76 s) /i:/ is faster than incorrect response (5.31 s), a difference of 1.55 s). Although the difference is very small (0.06 s), but the correct response (3.25 s) for /ʌ/ recorded a faster reaction time than the incorrect response (3.31 s). There is a huge reaction time different between the correct response and incorrect response of /ɑ:/. The correct response recorded a time of 3.21 s, while the incorrect response is at 5.33 s, a different of 2.12 s. Although the different is small (0.72 s), the participants' reaction time for the correct response of /ʊ/ is at 2.81 s, contrast to the incorrect response at 3.53

s. No data available to compare /u:/ since no incorrect response was recorded.

Table 4.11 Reaction Time to Response

Vowel	Correct Response	Total Reaction Time	Reaction Time to Response	Incorrect Response	Total Reaction Time	Reaction Time to Response
/ɪ/	18	92.1	5.12	12	64.96	5.41

/i:/	14	52.63	3.76	16	85	5.31
/ʊ/	29	81.42	2.81	1	3.53	3.53
/u:/	30	86.04	2.87	0	0	na
/ʌ/	21	68.18	3.25	9	29.78	3.31
/ɑ:/	21	67.35	3.21	9	48.01	5.33

Figure 4.27 shows the number of responses (both correct and incorrect responses) and the time taken by the participants to react (s). The trendline suggests that that the more correct response recorded, the shorter the time taken to react, while the more incorrect response chosen, the longer time taken by the participants to react.

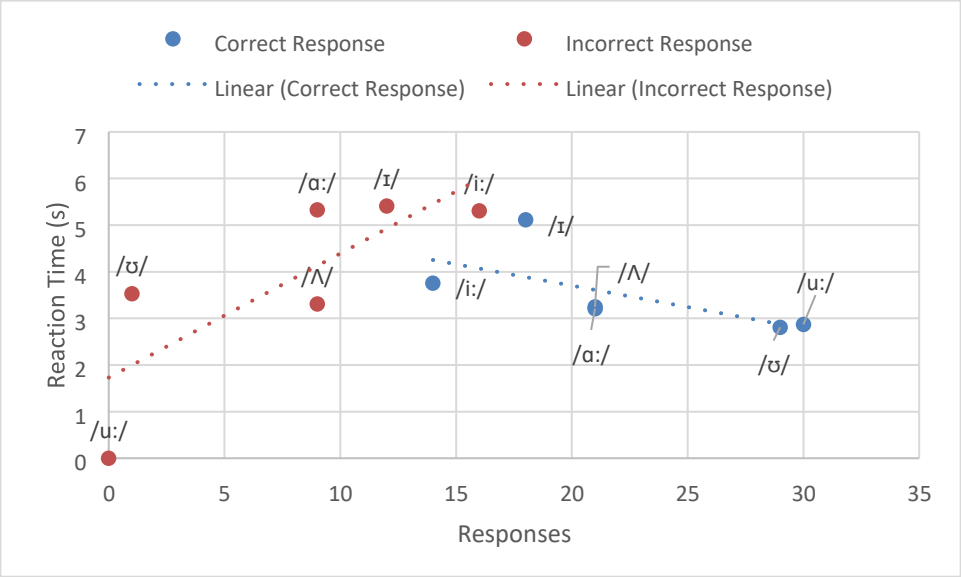

Figure 4.27 Relationship of Accuracy and Latency (by vowel) Figure 4.28 illustrates the relationship between the accuracy and latency based on vowel pair. Similar to Figure 4.30, the trendline also suggests that shorter times were recorded to produce more correct responses and vice versa. There were 32 correct responses of /ɪ/ - /i:/ compared to 28 incorrect responses, and the participants took 4.52 s to respond correctly and 5.36 s to respond incorrectly. Similarly, there were 42 correct responses for /ʌ/ - /ɑ:/, in which they took 3.23 s to respond while they took 3.53 s to respond to 18 incorrect responses. Likewise, the /ʊ/ - /u:/ also produced identical result as there was 59 correct response to one incorrect response, and the participants took 2.83 s to react to

correct response while 3.53 s to the incorrect response. Putting them in a hierarchy, /ʊ/ - /u:/ recorded the highest correct response and yet the participants took the least time to respond, while the /ɪ/ - /i:/ recorded the lowest correct response with the highest time taken to react. As for the incorrect response, /ɪ/ - /i:/ recorded the highest responses and time taken to react, in contrast with the /ʊ/ - /u:/, in which the pair recorded the lowest incorrect response with the lowest reaction time.

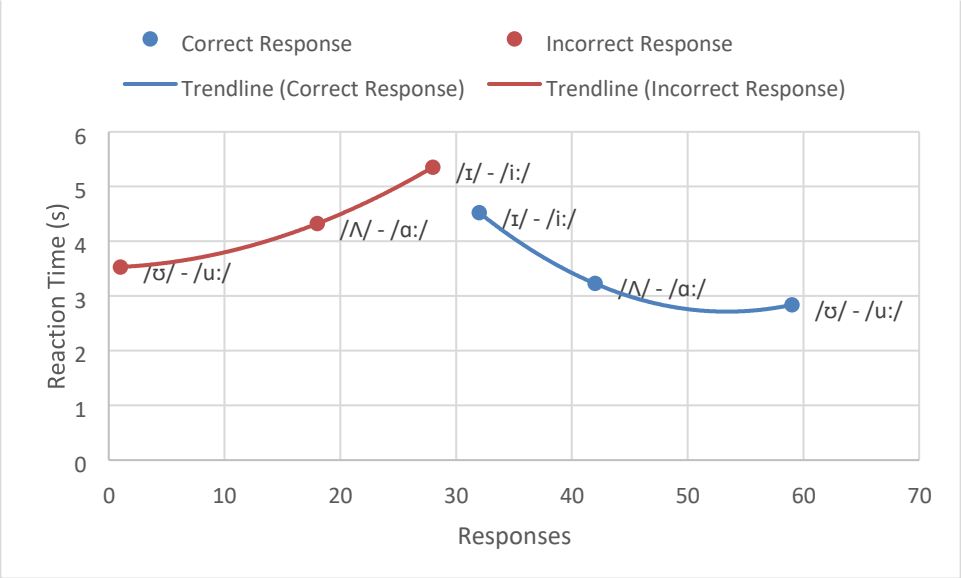

Figure 4.28 Relationship of Accuracy and Latency (by vowel pair)

After analysing the result, the vowel pair can be ranked based on the easiest to perceive. The pair /ʊ/ - /u:/ is the easiest to perceive (ranked 3rd), both in terms of accuracy and latency. The pair /ʌ/- /ɑ:/ is ranked second and follows by the pair /ɪ/ - /i:/ (the most difficult to perceive), as illustrated in Table 4.12.

Table 4.12 Result of Perception Test

Vowel		/ɪ/ - /i:/	/ʊ/ - /u:/	/ʌ/ - /ɑ:/
Correct Identification	Correct response	32	59	42
	Incorrect response	28	1	18
	Difference	4	58	24
Average Reaction Time	Correct response	4.52	2.84	3.40
	Incorrect response	5.55	3.53	3.93

(s)	Difference	1.03	0.69	0.53
Rank in Perception Level		1	3	2

**The pairs are ranked from the most difficult to clearly perceive (1) to the easier to clearly perceive (3).*

4.3.3 Relationship between Production and Perception of English Vowels

Table 4.13 presents the findings of the two tests, the production test and the perception test. Production test investigated the production of targeted vowels, in terms of vowel quality (F1 and F2) and duration, whereas the perception test investigated the perception of the targeted vowels in terms of accuracy (correct identification) and latency (reaction time). The data was gathered from 10 female Year 5 pupils.

/ɪ/ is produced at an F1 value of 4.733 Bark and an F2 value is 13.278 Bark, and its duration is 103.54 msec. The participants were able to correctly identify this vowel at 60% accuracy, in 3.07 s, while producing 40% inaccuracy in 2.17 s. The long vowel /i:/ is produced at F1 values of 4.621 Bark and F2 value of 12.757 Bark, with a longer duration of 107.87 msec. The participants poorly perform with only 46.67% accuracy at 1.75 s, and 53.33% inaccuracy at 2.83 s. In contrast, the participants perfectly identified /u:/ at 100% accuracy and took 2.87 s to respond. This vowel is produced at an F1 value of 4.947 Bark and an F2 value at 9.376 Bark, with a duration of 111.36 msec. As for /ʊ/, it is produced at an F1 value of 4.934 Bark and an F2 value of 8.686 Bark, with a duration of 108.16 msec. The perception test was almost perfectly scored with 96.67% accuracy at 2.81 s and 3.33% inaccuracy at 0.12 s. Lastly, the participants scored the same result in the accuracy test for both /Λ/ and /ɑ:/, at 70% accurate response and 30% inaccurate response. The only difference was on the latency, as /Λ/ took 2.27 s to react accurately and 0.99 s for an inaccurate while /ɑ:/ took 2.25 s for an accurate response and 1.60 s for an inaccurate response. As for the production, /Λ/ is produced at F1 value of 7.373 Bark

and F2 value at 11.435 Bark (duration of 108.23 msec) while /ɑ:/ is produced at F1 value of 7.339 Bark and F2 value at 11.238 Bark (157.23 msec).

Table 4.13 Findings of Production Test and Perception Test

Vowel	Production Test			Perception Test			
	F1	F2	Duration	Percentage of Correct Response	Percentage of Incorrect Response	Correct Response Reaction Time	Incorrect Response Reaction Time
/ɪ/	4.733	13.278	103.54	60	40	3.07	2.17
/i:/	4.621	12.757	107.87	46.67	53.33	1.75	2.83
/ʊ/	4.934	8.686	108.16	96.67	3.33	2.81	0.12
/u:/	4.947	9.376	111.36	100	na	2.87	Na
/ʌ/	7.373	11.435	108.23	70	30	2.27	0.99
/ɑ:/	7.339	11.238	157.23	70	30	2.25	1.6

The ability to perceive comes first in order to produce the sound, as suggested in the perception-production link. PAM suggested four categories to explain the perception of the L2, and this research supports two out of the four categories. SLM claims that there is a correspondence between the ability to perceive and to produce a language. Excitedly, the findings confirm this correspondence.

/ʌ/ and /ɑ:/ is the example of PAM Category Goodness because both sounds are assimilated into Malay /a/, but one is more deviant than the other. The participants scored both /ʌ/ - /ɑ:/ excellently, with 70% accuracy. Moreover, they can identify the long /ɑ:/ faster than the short /ʌ/. As discuss above, some teachers taught their students that words that are spelt with ‘ar’ will sound longer. This is in a way a good tip, hence the excellent performance of the participants when producing the pair. The vowel contrast may not be distinctly discriminated but the participants produced a wide contrast on the duration. The /ʊ/ - /u:/ pair is ranked as the easiest to both in production and perception. This pair falls under Category Goodness, as they can be perceived almost perfectly. The participants scored almost perfectly with 100% accuracy for /u:/ and 96.67% accuracy for /ʊ/. The

pair was also produced clearly acoustically and better duration contrast. One interesting finding on the vowel contrast of /ʊ/ and /u:/ pair is that, if the participants' production were to compare with the English vowel cardinal chart, the F2 position of /ʊ/ and /u:/ were actually incorrect as their places were swapped. However, since this research used Deterding (1997) mapping, the positions are correct. From this finding, it can be justified that even native speakers produced this sound incoherently. In addition, as found in Salwani (2005), both hers and this research participants did not contrast the pair in distinction. Moreover, her findings showed that some of the participants produced the short /ʊ/ longer than the long counterpart /u:/.

Similar situation occurred in this research but for the /ɪ/ - /i:/ pair. /ɪ/ and /i:/ can be under Single Category as these two vowels were assimilated into Malay /ɪ/. The findings showed that the participants identified the short /ɪ/ correctly by 60% and long /i:/ incorrectly (by selecting long /i:/ as short /ɪ/) by 53.33%. cross-reference to the production test, the participants did not manage to produce them acoustically accurate as the F2 position of the pairs were incorrect, as well as small duration contrast. SLM explained that if the L2 sound is perceived as similar to the L1 sound, thus they are assimilated, produce no difference between L2 and L1 sounds. In the present study, /ɪ/ - /i:/ pair is ranked as the most difficult to perceive, and also to produce. The Malay young speakers were having a hard time in producing and perceiving the /ɪ/ and /i:/ pair. This finding is consistent with other researches such as Perwitasari et al. (2016), Pillai & Delavari (2012) and Salwani (2005). The perception-production link suggested that the ability to perceive sounds correctly will influence the ability to produce that particular sounds properly. Many pieces of research which study the relationship of language perception and production both directly and indirectly supported this framework. Although some studies revealed that speaker who scores better in the production of the sound may not be able to perceive the same sound accurately. Research by Shahidi et al. (2012) found that the

participants were not able to perceive the target sounds accurately, even if they managed to produce them properly. Based on the perception-production link, the learners must acquire the perception of the sounds before being able to produce them accordingly. Referring back to Chapter 3, this research employed the production test first before the perception test. Contrary to Shahidi's findings, this research's participants still managed to perform properly. The reason for the production test conducted first is to ensure the result to be as authentic as possible. Since their perception is yet to be invoked, the participants were behaving and speaking as they normally do.

The participants' level of English proficiency may also influence their perception and production of English monophthong, as suggested in Ammar et al. (2016). Several of the participants claimed that they were exposed to the English language as early as three years old. Based on the Language Profile Test (refer to Table 3.4), a few of them rated themselves to be good and comfortable to speak and listen in English. In the production of vowel duration, P4 and P5 were having difficulties in producing the correct duration, as some short vowels were produced longer than the long vowel. This reflected their selfrated speaking fluency as they rated themselves with 1. As for P1 and P7, they self-rated themselves as 4 and 2 as they managed to produce short vowels shorter than the long vowels. Similarly, P1 and P7 did very well in the perception test, and this actually reflected their confidence and fluency of the language, as they did rate themselves quite high on the listening fluency (rated 1 and 2). Contrary, P4 and P5 only rated themselves 1 for listening fluency and they did not score much on the perception test. The participants of this research were undergone the previous version of KSSR, before the introduction of CEFR (refer to 2.1.3). Based on the syllabus, the content standard did emphasize language perception and language production. The revised version of the KSSR put more highlight on phonemic awareness than the previous version, as well as emphasizing language perception first, then language production. This approach is aligned with the

Perceptionproduction Link. More importantly, the decision to incorporate CEFR into the curriculum is correct, and many teachers are positively inclined as well as in favour of the CEFRaligned assessment (Azli & Aini, 2019). Looking at the finding, a good outcome on the production and perception of English sounds can be predicted in the future. In general, this research supports the perception-production link as when the speakers can perceive the sounds accurately, they can also produce the sound appropriately.

4.4 Conclusion

In conclusion, the participants of this research can successfully produce and perceive English short and long vowels. Despite the success, not all participants can perceive and produce each vowel accurately. /ɪ/ and /i:/ are the most difficult vowel pair to be perceived and produced, as presented in this research. Other than that, the participants can produce the pair /ʌ/ and /ɑ:/ correctly in terms of vowel contrast, and greatly in terms of duration. As for perception, the participants can discriminate between the two vowels. Lastly, /ʊ/ and /u:/ can be perceived accurately, although the production may not be accurately produced.

CHAPTER 5: CONCLUSION

The purpose of the present study is to explore how Malay female young learners of English produce and distinguish English monophthongs. Thus, this study examined the production of English short and long vowels, by looking at both formant frequencies and the duration of each vowel. In order to gain more insights, the research also investigated

the participants' perception of the English monophthongs by their ability to distinguish between short and long vowels. This chapter provides a summary of the finding, as well as discuss the result further based on the research question.

5.1 Summary of the Findings

10 Malay female young speakers of English participated in this study. The study was divided into two parts: production test and perception test. The production test provided data for Research Question 1 and 2, and the perception test prepared data for Research Question 3.

5.1.1 Research Question 1: What are the qualities of the vowel produced by Malay young learners of English in terms of formant frequencies (F1 and F2)?

The result of the production experiment shows that the participants were not aware of English vowel quality, as they produced English short and long vowels dissimilar to native speakers. The production of each vowel was mapped according to their F1 and F2 values (in Bark). The plotting revealed that the space of these participants is smaller and more to the back, in contrast to English native speakers. This finding is in agreement with other researches on L2 speakers' vowel space to be more compact than native counterparts (Perwitasari et al., 2016; Pillai & Jayapalan, 2010; Tan & Low, 2010).

The final results indicate that the F1 values of each vowel pair are within the same range, with very little difference of less than 0.1 Bark. As discussed in Chapter 4, the participants assimilated the L2 sounds into their L1 category. The pair /ʌ/ and /ɑ:/ were produced very closely, almost overlapping for both F1 and F2 values. Other than that, /ɪ/ - /i:/ were also exchanged their F2 position, as the result of many participants produced /ɪ/ as /i:/and vice versa. However, the production of /ʊ/ seems to be on the right track as it was produced more back compared to /u:/, showing a clear different value of F2. To put

the pair in order, /ɪ/ - /i:/ pair is the most difficult to produce, followed by /ʌ/ and /ɑ:/ and /ʊ/ - /u:/ pair.

5.1.2 Research Question 2: To what extent is the difference in duration between /ɪ/ - /i:/, /ʊ/ - /u:/, and /ʌ/ - /ɑ:/ vowels produced by Malay young learners of English?

The findings showed that the participants have the knowledge of vowel duration and all vowel pairs were successfully discriminated against. However, in this research, not all pairs were clearly distinguished. This is an interesting finding as it is different from what was found in previous researches. For example, Salwani (2005) in her research found that her participants were not able to clearly distinguish vowel duration, and her participants were young adults who majoring in the English language. Aside from the /ʌ/ - /ɑ:/ pair, this research participants did not clearly discriminate the duration between the short and long vowels. It was found that only 40% of the participants can discriminate the /ɪ/ from /i:/. The mean difference was between 0 - 29.0 ms, too short to be noticed by normal human hearing. Similarly, only 50% of the participants can discriminate between /ʊ/ and /u:/, with the mean difference of 0.3 – 10.24 ms. Again, the difference is too short. On contrary, all participants can discriminate /ʌ/ from /ɑ:/, with a mean different range from 0.79 to 80.86 ms. Only one participant barely discriminated the duration, while the rest of the participants produced quite a distinct difference in duration.

In the Malay language, the monophthong vowels consist of only single vowels with no indication of duration, hence the idea of vowels with different duration does not exist among Malay speakers. The new syllabus did well in providing awareness of English sounds, thus the improvement on the ability to distinguish the duration. However, it may take several years before the students may really be aware as well as understand English sound systems. As for /ʌ/ - /ɑ:/ pair, the distinction between these two vowels' duration is clearer. This unique finding resulting in the orthographic cue, the spelling of the stimuli,

which helped the participants to properly produce the sounds. The words used for /ɑ:/ are ‘cart’ and ‘dark’. Malay speakers, especially beginner learners of English, have the tendency to pronounce every letter of the word, as to how they do when speaking in their L1. When they encountered a word with 'r', Malay speakers tend to pronounce the /r/, as to how American English does. /r/ is an approximant and also considered as semi-vowel, as the sound is produced almost vowel-like. Subconsciously, the /ɑ:/ is produced with a distinct duration. To put the pair in order, /ʌ/ - /ɑ:/ pair is the easiest to produce, followed by /ɪ/ - /i:/ and /ʊ/ - /u:/pair.

5.1.3 Research Question 3: What is the ability of the Malay young learners in distinguishing English short and long vowels?

In general, the female Malay young learners were able to distinguish English short and long vowels, and they select the correct sounds faster than incorrect selections. The participants of this research have the ability to distinguish the /ʊ/ - /u:/ pair better than /ʌ/ - /ɑ:/ and /ɪ/ - /i:/ pairs. /ʊ/ - /u:/ pair has the highest accuracy percentage while /ɪ/ - /i:/ pair is the lowest. The participants managed to achieve more than 50% correct identification. The higher percentage of correct identification is on the /ʊ/ and /u:/, of 98.3%, follows by /ʌ/ and /ɑ:/ with 70%. Not surprisingly, the participants struggle to correctly identify between /ɪ/ and /i:/ with 53.3% correct identification.

Similar findings were found in terms of latency; /ʊ/ - /u:/ pair has the shortest reaction time towards the correct answer, compared to /ʌ/ - /ɑ:/ and /ɪ/ - /i:/ pairs. Mapping the latency of correct answer to the latency of the wrong answer, the results showed that the /ʊ/ - /u:/ pair is still produced a better result than /ʌ/ - /ɑ:/ and /ɪ/ - /i:/ pairs. As suggested in response latency by Schiepers (1980), inaccurate responses take a longer time to compute and to be selected. The finding in this research supported the statement, as the participants spent a shorter time to produce the correct response in contrast to the incorrect

response. Figure 4.30 and 4.31 illustrated this clearly. Each vowel pair can be categorised into one of the categories in PAM. The participants were not able to identify /ɪ/ - /i:/ clearly and this agrees with PAM Single Category and SLM. The participants took the longest time to respond and also produced the least correct answers. /ʌ/ - /ɑ:/ and ʊ/ - /u:/ pair belong to PAM Category Goodness. The participants were able to distinguish between the short and long vowels and respond within an appropriate time. As for /ʊ/ - /u:/, the participants were able to correctly identify the pair (the highest accuracy), in the shortest time (the highest latency). Thus, it is wise to rank the /ʊ/ - /u:/ pair as the most distinguishable, follow by /ʌ/ - /ɑ:/ and /ɪ/ - /i:/ pairs.

5.1.4 Discussion on Perception and Production of English Vowels

The perception-production link claims that the ability to perceive a sound influences the ability to produce it. SLM does support this claim as the ability to perceive a sound properly increases the ability to produce the sound correctly. Table 4.10 and Table 4.12 in Chapter 4 show the result for both the production and perception test. The \ʊ\ - \u:\ pair scores the highest rank for the production test. This pair was the only one with significant differences in vowel quality, though the difference in duration was not very strong. The participants can clearly perceive and distinguish the pair by correctly choosing the answers in the shortest time. On contrary, the pair \ɪ\ - \i:\ is the most difficult to produce and perceive. The pair has no significant difference in vowel quality and duration, score more incorrect responses with the longest time reaction. Lastly, the ability to perceive the /ʌ/-/ɑ:/ pair did assist them to produce a distinctly different duration, as illustrated in the production experiment. The result of this research support and is consistent with the PAM, SLM and perception-production link.

Table 5.1 Comparison between Vowel Pairs on the Result of Production and Perception Test

Vowel	F1 (Bark)	F2 (Bark)	Duration (s)	Accuracy	Latency
/ɪ/-/i:/	No different	Slight different	Medium different	Lowest	Lowest
/ʊ/-/u:/	Slight different	Major different	Minor different	Highest	Highest
/ʌ/-/ɑ:/	Slight different	Medium different	Major different	Medium	Medium

Table 5.1 proves the perception-production link theory. Bradlow et al. (1999), Isbell (2016) and Saito & Poeteren (2018) suggested that there is a tie between the perception and the production of sound among second language learners. The participants of the current research can perceive and distinguish /ʊ/ - /u:/ the best, and the understanding is clearly demonstrated as they can produce the pair better. As predicted, they performed badly in perceiving the /ɪ/ and /i:/, resulting in the poor sound production of the targeted pair. This result is consistent with previous research, speakers showed contrast in the production of the vowel pair when they can discriminate the contrast perceptually (Baker & Trofimovich, 2006; Yu, 2019).

5.2 Implication

This research provides an insight into understanding the production and perception of Malay speakers, especially among young female speakers. These findings add more input to the current body of information on English pronunciation by ESL learners. Based on the study, the area of Malay young speakers producing and perceiving English vowel sounds is still lacking in research, hence this opens up a potential research area, especially

now, with the implementation of KSSR and the introduction of CEFR recently. Indirectly, this research also provides an insight into how the new syllabus, especially on English, could improve the students' performance, in this case, in terms of speaking (pronunciation). This research provides an insight into understanding the production and perception of Malay speakers, especially among young female speakers. This research also focuses on three vowel pairs, which allows researchers to understand the vowel space produced by the participants. After all, different from previous researches who investigated young adult and adult speakers, this study opens up a new perspective and potential research area, by looking at how young learners in their primary school produce and perceive English vowels sounds. As mentioned in 1.6, there are limitations to this study, hence the finding cannot be generalized to other populations, but the methods can be adapted. Thus, a few suggestions and recommendations are proposed for future researches.

5.3 Recommendations

This research has resulted in some fruitful findings regarding the production and perception of English short and long vowels in terms of vowel frequency, duration, accuracy and latency. At the same time, it also reveals a few issues that can be looked into and provides improvement and filling up gaps.

It is recommended to future researchers to go deeper into these aspects:

i. Race and first language: Other than gender and age, future research can also investigate how the race of the participant may contribute to the outcome of the research. A better understanding of whether the first language interference does affect the outcome. Many past pieces of research focused on Malay speakers, and various studies also only study one race at a time. A comparison between speaker

races and first language could widen the research area, as well as provide a better insight on second language learners.

ii. Gender: Future studies can instead employ male participants rather than female participants. There was only a handful of research that looked at male participants compared to female participants. Moreover, it is understood that male voice production is different from the female, hence it could result in different and interesting findings.

iii. Stimuli: Since this research only focuses on and /ɪ/, /iː/, /ʌ/, /ɑː/, /ʊ/ and /uː/, future research may look at different monophthong such as /ə, /ɜː/, /ɔː/, /ɒ/ and /e/. the research can either study the vowel pairs, or the vowel individually. Aside from monophthong vowels, the researcher can use a different stimulus such as diphthongs: /aɪ/, /eɪ/, /əʊ/, /aʊ/, /eə/, /ɪə/ and /ɔɪ/. and triphthong, or even looking at the consonants.

iv. Population: It is recommended for future study to look into speakers from rural school as their exposure to the English language are increasing, thus this could bring new insight into this area.

5.4 Conclusion

In conclusion, this research has provided information and explanation on the production and perception of English short and long vowels among Malay young female speakers. This research investigated whether they could produce the vowel sound correctly in terms of frequency and duration. The findings showed that the speakers may not be able to clearly produce the sound at the appropriate frequency but can somehow maintain a good duration difference between the short and long vowels. The participants' perception of English monophthong was evaluated based on accuracy in the identification test and the time taken to correctly select the answers. As the result, the speakers were

having a problem with /ɪ/ - /i:/ pair but better with /ʊ/ - /u:/ pair. Interestingly, they spent less time choosing correct answers but more time when selecting the incorrect answer.

It is hoped that this study may help many, not only future researchers but also teachers who are teaching young learners English as their second language. With the understanding of how these participants produced and perceived English, teachers can plan a better teaching approach in teaching English, especially in reading and pronunciation. The findings may not be generalized to other populations but can be used as guidelines and references to assist understanding.

REFERENCES

Ahmadi, M. R., & Abbas, P. G. (2011). Why is pronunciation so difficult to learn? *English Language Teaching*, *4*(3), 74–83. https://doi.org/10.5539/elt.v4n3p74

Ammar, A.-W. A. A., Ilyana, J., Che, A. A. G., & Yap, N. T. (2016). First language and proficiency level effects on English vowel perception by Iraqi learners of English in Malaysia. *ASEAN Journal of Teaching and Learning in Higher Education (AJTLHE)*, *8*(1), 26–45.

Amutan, K. I., Ramalingam, S. J., Hee, S. C., & Maruthai, E. (2017). The selection practices and recruitments of fresh graduates in local organisation's job interview. *Journal of Language and Communication*, *4*(2), 153–167.

Azli, N. A. L., & Aini, A. M. K. (2019). Implementation of CEFR-aligned assessment tools in Malaysian ESL classroom. In *Asia Proceeding of Social Sciences* (Vol. 4, pp. 7–10). Islamabad: Reader Insight Publisher.

Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics*, *44*, 231–250. https://doi.org/10.1515/IRAL.2006.010

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: the role of cross-language similarity. *Language and Speech*, *51*(4), 317–342. https://doi.org/10.1177/0023830908099068

Bello, H., Ngee, T. Y., Mei, Y. C., & Nimehchisalem, V. (2020). An acoustic analysis of English vowels produced by Nigerian and Malaysians ESL speakers. *Journal of Language and Communication*, *7*(1). https://doi.org/10.24303/lakdoi.2015.23.1.65

Best, C. T. (1991). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The Development of Speech Perception*, 1–30. https://doi.org/10.7551/mitpress/2387.003.0011

Best, C. T., Goldstein, L., Tyler, M. D., & Nam, H. (2009). Articulating the Perceptual Assimilation Model (PAM): Perceptual assimilation in relation to articulatory organs and their constriction gestures. *The Journal of the Acoustical Society of America*, *125*(4), 2758–2758. https://doi.org/10.1121/1.4784648

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. . Munro & O. S. Bohn (Eds.), *Language experience in second language speech learning* (pp. 13–34). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.17.07bes

Best, C. T., Tyler, M. D., Goldstein, L., & Nam, H. (2016). Articulating what infants attune to in native speech. *Ecological Psycholo*, *28*(4), 216–261. https://doi.org/10.1080/10407413.2016.1230372

Birdsong, D., Gertken, L. ., & Amengual, M. (2012). Bilingual Language Profile: An easy-to-use instrument to assess bilingualism. Retrieved from https://sites.la.utexas.edu/bilingual/

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer. Retrieved from www.praat.org

Bradlow, A. R., Akahane-yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perceptual Psychophysics*, *61*(5), 977–985.

Burman, D. D., Bitan, T., & Booth, J. R. (2008). Sex differences in neural processing of language among children. *Neuropsychologia*, *46*(5), 1349–1362. https://doi.org/10.1016/j.neuropsychologia.2007.12.021

Casillas, J. (2015). Production and perception of the /i/-/I/ vowel contrast: The case of L2-Dominant early learners of English. *Phonetica*, *72*(2–3), 182–205. https://doi.org/10.1159/000431101

Casillas, J. V, & Simonet, M. (2016). Production and perception of the English / æ /–/ ɑ / contrast in switched-dominance speakers. *Second Language Research*, *32*(2), 171–195. https://doi.org/10.1177/0267658315608912

Clynes, A., & Deterding, D. (2011). Standard Malay (Brunei). *Journal of the International Phonetic Association*, *41*(2), 259–268. https://doi.org/10.1017/S002510031100017X

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, *27*, 47–55. https://doi.org/10.1017/S0025100300005417

Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy: A 10- to 12second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology*, *87*(5), 951–962. https://doi.org/10.1037/0021-9010.87.5.951

Elvin, J., Williams, D., & Escudero, P. (2016). The relationship between perception and production of Brazilian Portuguese vowels in European Spanish monolinguals. *Loquens*, *3*(July 2016), 2386–2637. https://doi.org/http://oadoi.org/10.3989/loquens.2016.031

Elya, D. T., & Parilah, M. S. (2015). Teacher's knowledge of phonemic awareness and its instruction in ESL learning. *International Journal of Technical Research and Applications*, (22), 72–79.

Employers : Fresh graduates have unrealistic expectations. (2015). Retrieved June 6, 2016, from http://www.jobstreet.com.my/career-resources/employers-freshgraduates-unrealistic-expectations/#.V1d66TV97Dc

English Language Standards and Quality Council. (2015). *English language education reform in Malaysia - The roadmap 2015-2025*. (M. D. Zuraidah, Ed.). Ministry of Education Malaysia.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(04). https://doi.org/10.1017/S0272263104040021

Faiz, H. (2018, October 5). poor command of English—reason graduates fail at job interviews. *UPM News*. Retrieved from https://sgs.upm.edu.my/news/poor_command_of_englishreason_graduates_fail_at_job_interviews-45281

Flege, J. E. (n.d.). SLM Speech Learning Model. Retrieved September 20, 2018, from http://www.jimflege.com/discussion_Topic1.html

Flege, J. E. (1995a). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in CrossLanguage Research* (p. 233). York Press.

Flege, J. E. (1995b). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics*, *16*(4), 425–442.

https://doi.org/10.1017/s0142716400066029

Flege, J. E. (2007). Language contact in bilingualism: Phonetics system interactions. In J. Cole & J. I. Hualde (Eds.), *Laboratory Phonology 9* (pp. 353–381). New: Mouton de Gruyter. https://doi.org/10.1515/9783110851847

Gurnam, K. S., & Suthagar, N. (2010). *Say it right*. Shah Alam, Selangor: University Publication Centre (UPENA) UiTM.

Hawkins, S., & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, *35*(2), 183–199. https://doi.org/10.1017/S0025100305002124

Hubais, A., & Pillai, S. (2010). An instrumental analysis of English vowels produced by Omanis. *Journal of Modern Languages*, *20*, 1–18.

IBM Corp. (2010). IBM SPSS Statistics for Window. Armonk, New York: IBM Corp. Retrieved from www.ibm.com/

Idham, M. M. R., Yusof, A. M., Syazana, W. N., Effa, W. J., & Talib, A. H. (2014). Factors influencing unemployment among graduates in Malaysia – An overview. *Journal of Economics and Sustainable Development*, *5*(11). https://doi.org/10.1016/j.sbspro.2014.01.1269

IELTS. (n.d.). How are your IELTS Speaking scores calculated? | IDP IELTS. Retrieved October 25, 2020, from https://www.ielts.com/malaysia/results/scores/speaking

Isbell, D. (2016). The perception-production link in L2 phonology. *MSU Working Papers in Second Language Studies*, *7*(October 2016), 57–67.

Jacewicz, E., Fox, R. A., & Salmons, J. (2011). Vowel change across three age groups of speakers in three regional varieties of American English. *Journal of Phonetics*, *39*(4), 683–693. https://doi.org/10.1016/j.wocn.2011.07.003

Jayapalan, K., & Pillai, S. (2011). The state of teaching and learning English pronunciation in Malaysia: a preliminary study. *Malaysian Journal of ELT Research*, *7*(2), 63–81.

Kendall, T., & Fridland, V. (2012). Variation in perception and production of mid front vowels in the U . S . Southern Vowel shift. *Journal of Phonetics*, *40*(2), 289–306. https://doi.org/10.1016/j.wocn.2011.12.002

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*. San Diego: Singular Publishing Group.

Kleber, F., & Harrington, J. (2011). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech*, *55*(3), 383–405. https://doi.org/10.1177/0023830911422194

Komurcu, B., & Yildiz, Y. (2011). An assessment of the auditory speech perceptual skills of young Turkish learners of English: Perceptual Assimilation Model. In *Procedia - Social and Behavioral Sciences* (Vol. 15, pp. 3511–3515). Elsevier B.V. https://doi.org/10.1016/j.sbspro.2011.04.327

Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *The Journal of the Acoustical Society of America*, *124*(6), 3959–3971. https://doi.org/10.1121/1.2999341

Lim, S. K., & Nazri, M. latiff A. (2013). An investigation of phonetics instruction in acquiring English pronunciation amongst Polytechnic learners using questionnaire. *Journal of Teaching and Education*, *2*(4), 257–269. Retrieved from http://www.universitypublications.net/jte/0204/html/P3G183.xml

Malaysia Education Blueprint 2013 - 2025. (2013). *Education*, *27*(1), 1–268. https://doi.org/10.1016/j.tate.2010.08.007

Malaysian Examinations Council. (2016). *Malaysian University English Test (MUET)*. *Batu Caves: Percetakan Warni*. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Malaysian+university+english+test#4

Malaysian graduates face job interview jitters, no thanks to weak English skills. (2015, November 18). *Malaysian Insider*. Retrieved from https://www.todayonline.com/world/asia/malaysian-graduates-face-job-interviewjitters-no-thanks-weak-english-skills

Merriam-Webster. (n.d.). Minimal Pair. Retrieved December 26, 2020, from https://www.merriam-webster.com/dictionary/minimal pair

Mulligan, K., Grant, J. T., Mockabee, S. T., & Monson, J. Q. (2003). Response Latency Methodology for Survey Research: Measurement and Modeling Strategies. *Political Analysis*, *11*(3), 289–301. https://doi.org/10.1093/pan/mpg004

Nasrin, M. A. R. I. (2018, October 5). Tidak fasih berbahasa Inggeris antara punca gagal mendapat pekerjaan. *Berita Harian*.

Ng, A. W. Y., & Chan, A. H. S. (2012). Finger response times to visual, auditory and tactile modality stimuli. In *International MultiConference of Engineers and Computer Scientists* (Vol. II). Hong Kong.

Nik, S. K., Farid, M. O., Hashim, M., & Hamid, A. M. (2008). *Tatabahasa dewan* (3rd ed.). Kuala Lumpur: Dewan Bahasa dan Pustaka.

Norsimah, M. A., Kesumawati, A. B., Norzakiah, A. H., & Norhashimah, J. (2007). Morphological differences between Bahasa Melayu and English: Constraints in students' understanding. In *The Second Biennial International Conference on Teaching and Learning of English in Asia : Exploring New Frontiers (TELiA2)* (pp. 1–11). Holiday Villa Beach & Spa Resort, Langkawi. Retrieved from http://repo.uum.edu.my/3264/1/N6.pdf

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*(May 2014). https://doi.org/10.3389/neuro.11.010.2008

Peirce, J. W. (2017). Psychopy. Retrieved from www.psychopy.org

Perwitasari, A., Klamer, M., & Schiller, N. O. (2015). Vowel duration in English as a second language among Javanese learners. In *18th International Congress of Phonetic Sciences* (pp. 1–5). Glasgow: University of Glasgow: Glasgow. https://doi.org/10.13140/RG.2.1.1752.5520

Perwitasari, A., Klamer, M., & Schiller, N. O. (2016). Formant frequencies and vowel space area in Javanese and Sundanese English language learners. *3L The Southeast Asian Journal of English Language Studies*, *22*(3), 141–152. https://doi.org/10.17576/3L-2016-2203-10

Pillai, S. (2007). Speaking English the Malaysian way : Is there a standard Malaysian English pronunciation?

Pillai, S. (2014). *The monophthongs and diphthongs of Malaysian English: An instrumental analysis*. *Peter Lang*. Retrieved from http://thestar.com.my/lifestyle/story.asp?file=/2011/4/22/lifefocus/8507563&sec=lifefocus

Pillai, S. (2017). Local features of English pronunciation : to embrace or ignore in the ELT classroom ? *Journal of English Language, Literature, and Teaching*, *1*(January), 1–8.

Pillai, S., & Delavari, H. (2012). The production of English monophthong vowels by Iranian EFL learners. *Poznan Studies in Contemporary Linguistics*, *48*(3), 473–493. Retrieved from English monophthong vowels; English as a Foreign Language (ESL); Vow- el contrast; Persian.%0A1.

Pillai, S., Don, Z. M., Knowles, G., & Tang, J. (2010). Malaysian English : an instrumental analysis of vowel contrasts. *World Englishes*, *29*(2), 159–172.

Pillai, S., & Jayapalan, K. (2010). The teaching of pronunciation in Malaysia: State of the art or no state at all? In *Transformations in English Language Education Proceedings of the 19th MELTA International Conference*.

Prodanovska-Poposka, V. (2017). A study of proper pronunciation as a factor of successful communication. In *CBU International Conference of Innovations in Science and Education* (Vol. 5, pp. 778–783). Prague. https://doi.org/10.12955/cbup.v5.1024

Racine, J. P. (2013). Reaction time methodologies and lexical access in applied linguistics. *Vocabulary Learning and Instruction*, (Vl), 1–5. https://doi.org/10.7820/vli.v03.1.racine

Rajadurai, J. (2006). Pronunciation issues in non-native contexts: a Malaysian case study. *Malaysian Journal of ELT Research*, *2*(1), 18.

Roach, P. (2007). *English phonetics and phonology: A practical course* (3rd ed.). United Kingdom: Cambridge University Press.

Saito, K., & Poeteren, K. Van. (2018). *The perception-production link revisited : The case of Japanese learners ' English /r/ performance*. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12175

Salwani, W. A. W. A. (2005). *Vowel length discrimination among Malay speaker of Malaysian English: An instrumental study*. Kuala Lumpur: Fakulti Bahasa dan Linguistik Universiti Malaya.

Sani, R. (2019, March 6). Upskilling with English. *New Straits Times*. Retrieved from https://www.nst.com.my/education/2019/03/466461/upskilling-english

Schepens, J. J., van der Slik, F. W. P., & van Hout, R. W. N. M. (2015). The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLOS ONE*, *10*(11), e0142056. https://doi.org/10.1371/journal.pone.0142056

Schiepers, C. (1980). Response latency and accuracy in visual word recognition. *Perception & Psychophysics*, *27*(1), 71–81. https://doi.org/10.3758/BF03199908

Shahidi, A. H., Aman, R., & Kechot, A. S. (2012). Production and perception of English word final stops by Malay speakers. *GEMA Online Journal of Language Studies*, *12*(4), 1109–1125.

Shak, P., Lee, C. S., & Stephen, J. (2016). Pronunciation problems : A case study on English pronunciation errors of low proficient students. *International Journal of Language Education and Applied Linguistics*, *04*, 25–35.

Soozandehfar, S. M. A. (2012). How to Improve Pronunciation? An In-depth Contrastive Investigation of Sound-spelling Systems in English and Persian. *International Journal of Instructional Technology and Distance Learning*, *29*(2), 75–93. https://doi.org/10.4304/jltr.2.5.1086-1098

Tan, R. S. K., & Low, E.-L. (2010). How different are the monophthongs of Malay speakers of Malaysian and Singapore English? *English World-Wide*, *31*(2), 162–189. https://doi.org/10.1075/eww.31.2.03tan

Thai, Y. N., Eng, W. B., & Aziz, A. Y. A. (2010). Representation of English front vowels by Malay-English bilinguals. *Pertanika Journal of Social Science and Humanities*, *18*(2), 379–389.

Thomas, E. (2010). *Sociophonetics: An introduction*. Macmillan International Higher Education.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4–21. https://doi.org/10.1159/000356237

Why many graduates can't express themselves in English. (2017, October 1). *New Straits Times*. Retrieved from https://www.pressreader.com/malaysia/new-straitstimes/20171001/282591673144317

Yamaguchi, T., & Chiew, P. S. (2019). Is there conflation? An acoustic analysis of vowels in Japanese English. *Asian Englishes*, *00*(00), 1–17. https://doi.org/10.1080/13488678.2018.1563665

Yang, J., & Fox, R. A. (2014). Perception of English vowels by bilingual Chinese–English and corresponding monolingual listeners. *Language and Speech*, *57*(2), 215–237. https://doi.org/10.1177/0023830913502774

Yu, A. C. L. (2019). On the nature of the perception-production link: Individual variability in English sibilant-vowel coarticulation. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *10*(1), 1–29. https://doi.org/10.5334/labphon.97

Zahariah, P. (2005). Perception of voicing in English word-final obstruents by Malay speakers of English : Examining the Perceptual Assimilation Model. *Malaysian Journal of ELT Research*, *1*, 1–12.