THE PRODUCTION AND PERCEPTION OF ENGLISH MONOPHTHONGS BY ACEHNESE SPEAKERS

TANZIR MASYKAR

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITI MALAYA
KUALA LUMPUR

2023

THE PRODUCTION AND PERCEPTION OF ENGLISH MONOPHTHONGS BY ACEHNESE SPEAKERS

TANZIR MASYKAR

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITI MALAYA
KUALA LUMPUR

2023

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Tanzir Masykar

Matric No: THA160034/17006293/1

Name of Degree: DOCTOR OF PHILOSOPHY

Title of ~~Project Paper/Research Report/Dissertation~~/Thesis ("this Work"): THE PRODUCTION AND PERCEPTION OF ENGLISH MONOPHTHONGS BY ACEHNESE SPEAKERS

Field of Study: World Englishes

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: 26 August 2023

Subscribed and solemnly declared before,

Witness's Signature Date: 26 August 2023

Name:

Designation:

THE PRODUCTION AND PERCEPTION OF ENGLISH MONOPHTHONGS BY ACEHNESE SPEAKERS

ABSTRACT

English is the only foreign language made as a compulsory subject from junior to senior high school in the Indonesian school curriculum. Some private schools also promote the use of English outside school hours. One such school system is the Islamic boarding high schools in Aceh, which require students to speak in English with their peers, seniors and teachers during their three years stay at the school. The rationale for the use of English is to improve the English language fluency of the students. However, to date, there is a lack of published studies on the English sounds produced by the students. To begin filling this gap, the present study investigated the production and perception of English monophthongs by Acehnese Indonesian speakers from one Islamic boarding school in Aceh. A total of 29 Islamic boarding high school students participated in the study. Two types of data were collected: (i) the production of English, Acehnese, and Bahasa Indonesia monophthongs, and (ii) the perception of English vowel contrast. For the production data, students produced English, Acehnese and Bahasa Indonesia monophthongs embedded in target words placed in a carrier sentence. For the perception data, students completed the AX, ABX and FCI tests of English vowel contrast with the audio stimuli provided by an American English speaker. The findings revealed that the quality of English monophthongs produced by the speakers were similar to Acehnese Indonesian speakers reported in a previous study from public non-boarding schools. Second, the speakers tended not to contrast typical vowel pairs in terms of quality. Third, three types of assimilation were found, full assimilation, partial assimilation and mixed assimilation. Fourth, the perception accuracy of each vowel pair was found to be task dependent. Their discrimination score was mostly higher in the ABX, followed by the AX and FCI tests. Finally, their ability to discriminate English vowel pairs in the perception test was assisted not by their assimilation of English vowels to Acehnese and vowels as suggested by speech learning models such as Perceptual Assimilation Model (PAM) and Second Language Linguistic Perception (L2LP). Instead, the cues they heard from the audio stimuli assisted their ability to discriminate the vowel contrasts.

Keywords: Acehnese-Indonesian speakers, English monophthongs, production and perception, vowel contrast

PRODUKSI DAN PERSEPSI BUNYI MONOFTONG BAHASA INGGERIS OLEH PENUTUR ACEH

ABSTRAK

Bahasa Inggeris adalah satu-satunya bahasa asing yang dijadikan sebagai mata pelajaran wajib dari sekolah rendah hingga menengah atas dalam kurikulum sekolah Indonesia. Beberapa sekolah swasta juga menggalakkan penggunaan bahasa Inggeris di luar waktu sekolah. Salah satu sekolah tersebut ialah sekolah menengah berasrama Islam di Aceh yang memerlukan pelajar bertutur dalam bahasa Inggeris dengan rakan sebaya, senior dan guru selama tiga tahun berada di sekolah tersebut. Rasionalnya ialah penggunaan bahasa Inggeris yang berterusan meningkatkan kemahiran bahasa pelajar. Walau bagaimanapun sehingga hari ini masih kurang kajian yang diterbitkan berkaitan dengan bunyi-bunyi bahasa Inggeris yang dihasilkan oleh pelajar-pelajar ini. Bagi mengisi jurang ini, kajian ini menyelidik pengeluaran dan persepsi monoftong Inggeris oleh penutur Aceh Indonesia dari salah sebuah pondok pesantren di Aceh. Seramai 29 orang pelajar sekolah menengah pondok (15 lelaki dan 14 perempuan) telah mengambil bahagian dalam kajian ini. Dua jenis data telah dikumpul: (i) penghasilan monoftong bahasa Inggeris, Aceh, dan Indonesia, dan (ii) persepsi kontras monoftong Inggeris. Untuk data pengeluaran, pelajar menghasilkan monoftong bahasa Inggeris, Aceh dan Indonesia yang dimasukkan dalam perkataan sasaran yang diletakkan dalam ayat pembawa. Untuk data persepsi, pelajar mengambil ujian AX, ABX dan FCI kontras monoftong Inggeris dengan rangsangan audio yang disediakan oleh penutur bahasa Inggeris Amerika. Dapatan kajian menunjukkan bahawa kualiti monoftong bahasa Inggeris yang dihasilkan oleh penutur semasa adalah serupa dengan penutur Indonesia Aceh yang dilaporkan dalam kajian lepas dari sekolah bukan berasrama awam. Kedua, pelajar cenderung tidak membezakan pasangan bunyi vokal yang biasa dari segi kualiti bunyi. Ketiga, beberapa jenis asimilasi daripada L3 Inggeris kepada L1 Aceh dan L2 Indonesia ditemui, asimilasi penuh,

asimilasi separa dan asimilasi campuran. Keempat, ketepatan persepsi setiap pasangan vokal didapati bergantung kepada jenis ujian. Skor diskriminasi mereka kebanyakannya lebih tinggi dalam ABX, diikuti dengan ujian AX dan FCI. Akhirnya, keupayaan mereka untuk mendiskriminasi pasangan vokal bahasa Inggeris dalam ujian persepsi bukan dipengaruhi oleh asimilasi bunyi vokal bahasa Inggeris kepada vokal Aceh dan Indonesia sepertimana yang disarankan oleh *Perceptual Assimilation Model (PAM)* dan *Second Language Linguistic Perception (L2LP)*. Sebaliknya, isyarat yang mereka dengar daripada rangsangan audio membantu keupayaan mereka untuk mendiskriminasi kontras vokal.

Kata kunci: Penutur Bahasa Aceh-Indonesian, monoftong Bahasa Inggeris, produksi dan persepsi, kontrast vokal

ACKNOWLEDGEMENTS

Alhamdulillah, I am grateful to Almighty Allah for giving me the strength to finalize this thesis.

The completion of this thesis would not have been possible without the valuable guidance and enduring support from my supervisors at Universiti of Malaya. My sincere and deep sense of gratitude goes to Prof. Dr. Stefanie Shamila Pillai and Assoc. Prof. Dr. Roshidah Binti Hassan for sharing their valuable and constructive feedback throughout my PhD program. Their willingness to give their time so generously to supervise and motivate me throughout my study is very much appreciated.

This PhD was financially supported by the Indonesian Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan). Therefore, I am deeply indebted to the Indonesian Government for granting me the scholarship, giving me the opportunity to pursue my PhD at such a prestigious university.

I would also like to thank my beloved wife, Febri Nurrahmi who has supported, guided, and motivated me through the ups and down in my PhD journey, my son Muhammad Hanif Alfaatih, and daughter Khadijatul Hunaifaa,who have been patient and understanding when I have to work long hours and did not have much time to spend with them. I also would like to thank my sister Nurul Taflihati Masykar who helped me during the data collection process, Sabariah and Aguslijar who organized the students and helped me to get the permissions from the school. Special thanks go to Faisal Mustafa who helped shed light on my data analysis.

I would also like to extend my deepest appreciation to my beloved Father (Muhammad Abbas Ibrahim), mother (Kartini MS), mother-in-law (Tjut Mariam Zanaria), father-in-law (Teuku Alfiansyah), brother (Tabsyir Masykar), sister-in-law (Ade Oktivoyari and

Ayu Imanidar), and brother-in-law (Darmawan) for their prayer, support and encouragement throughout my PhD programme.

TABLE OF CONTENTS

THE PRODUCTION iii

ABSTRAK iv

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS viii

LIST OF FIGURES xiii

LIST OF TABLES xvi

LIST OF APPENDICES xix

LIST OF ABBREVIATION xx

CHAPTER 1: INTRODUCTION 1

1.1 Introduction 1

1.2 Languages in Indonesia 1

1.3 English in Islamic Boarding Schools 5

1.4 Problem Statement 7

1.5 Research Aim and Objectives 10

1.6 Research Questions 11

1.7 Significance of Study 11

1.8 Scope of Study 12

1.9 Thesis Outline 12

CHAPTER 2: LITERATURE REVIEW 14

2.1 Introduction 14

2.2 Global Englishes as a Theoretical Framework 14

2.2.1 Kachru's Three Circles Model 15

2.2.2 Schneider's Dynamic English Model 21

2.2.3 Jenkin's English as a Lingua Franca 25

2.3 Positioning English in Indonesia in relation to Neighbouring Countries. 27

2.3.1 English in Indonesia's Neighbouring Countries 27

2.3.2 English and the Regional Language in Indonesia 33

2.3.2.1 Bahasa Indonesia 38

2.3.2.2 Acehnese Language 42

2.3.2.3 Acehnese Monophthongs 44

2.3.2.4 Bahasa Indonesia Monophthongs 45

2.3.2.5 English Monophthongs 46

2.3.2.6 English vowel contrast 47

2.4 Studies on English Vowel Production by Indonesian Speakers 49

2.5 Models of Speech Perception 56

2.6 Studies on the English Vowel Perception by Indonesian Speakers 63

2.7 Conclusion 66

CHAPTER 3: RESEARCH METHODOLOGY 67

3.1 Introduction 67

3.2 Research Design 67

3.3 Participants 68

3.4 Data Collection 69

3.4.1 Production Test 70

Elicitation materials 72

3.4.2 Perception Task 74

ABX Task 76

AX Task 78

FCI Task 81

3.5 Ethical Considerations 82

3.6 Data Analysis 83
3.7 Vocabulary Size Test 89
3.8 Word Familiarity Index 91

CHAPTER 4: PRODUCTION OF THE ENGLISH MONOPHTHONGS BY ACEHNESE-INDONESIAN SPEAKERS 96

4.1 Introduction 96
4.2 Quality of English Monophthongs 96
4.2.1 Front Vowels 98
4.2.2 Mid Vowel 102
4.2.3 Back Vowels 105
4.3 Vowel Contrast 109
4.3.1 Vowel Contrast for Bit-Beat 110
4.3.2 Vowel Contrast for Bet and Bat 111
4.3.3 Vowel Contrast for Bud-Pot 112
4.3.4 Vowel Contrast for Pot-Port 114
4.3.5 Vowel Contrast for Foot-Food 115
4.3.6 Durational Contrast 116
4.4 Conclusion 118

CHAPTER 5: COMPARISON OF ENGLISH MONOPHTHONG AND EQUIVALENT ACEHNESE-INDONESIAN VOWELS 120

5.1 Introduction 120
5.2 Quality of Acehnese Monophthongs 120
5.3 Quality of Indonesian Monophthong 123
5.4 English Vowels Similar to Acehnese and or Indonesian 125
5.4.1 Vowel [i:] in beat 127

5.4.2 Vowel [ɔː] in port 129
5.4.3 Vowel [uː] in food 131
5.4.4 Vowel [ɛ] in bet 133
5.4.5 Vowel [ʌ] in bud 134
5.4.6 Vowel [ɪ] in bit 138
5.4.7 Vowel [æ] in bat 140
5.4.8 Vowel [ɑː] in pot 142
5.4.9 Vowel [ʊ] in foot 143
5.5 Conclusion 146

CHAPTER 6: PERCEPTION OF ENGLISH VOWELS 147

6.1 Introduction 147
6.1 A comparison of English, Acehnese, and Indonesian vowels 147
6.1.1 Findings from the AX Task 150
6.1.2 Findings from ABX Task 156
6.1.3 Findings from Forced Choice Identification (FCI) Task 161
6.1.4 Comparisons of the Three Tasks 166
6.2 Production and Perception Relationship 169
6.2.1 Comparison with the American Native Speaker of English 169
6.2.2 Perceptual Assimilation Model 177
6.2.3 Second Language Linguistic Perception Model 189
6.3 Conclusion 191

CHAPTER 7: DISCUSSION 193

7.1 Introduction 193
7.2 English monophthong quality 193
7.3 Quality Contrast 199

7.4 Length Contrast 205
7.5 English vowels quality to equivalent Acehnese and Indonesian vowels 208
7.6 Perception of Monophthong Contrasts 212
7.7 Perception and Production 215
7.8 Conclusion 219

CHAPTER 8: CONCLUSION 221
8.1 Conclusion 221
8.2 Limitation 222
8.3 Suggestions 224
8.4 Implications 225
REFERENCES 227
List of Publications and Papers Presented 240
LIST OF APPENDICES 241

LIST OF FIGURES

Figure 3.1: Onset and offset of English [iː] in beat 85

Figure 3.2: Number of speakers familiar with words used in the study 94

Figure 3.3: Comparison of word familiarity index in word pairs 95

Figure 4.1: Vowel plot of English monophthongs 98

Figure 4.2: Front vowels 99

Figure 4.3: Box plot of F1 and F2 by gender for English vowel [ɪ] in *bit* 100

Figure 4.4: Box plot of F1 and F2 by gender for English vowel [iː] in *beat* 101

Figure 4.5: Box plot of F1 and F2 by gender for English vowel [ɛ] in *bet* 101

Figure 4.6: Box plot of F1 and F2 by gender for English vowel [æ] 102

Figure 4.7: Mid vowels 103

Figure 4.8: Box plot of F1 and F2 by gender for English vowel [ʌ] in *bud* 103

Figure 4.9: Comparison of English [ʌ] to English [ʊ] and [uː] (left) and English [ʌ] with speaker labels (right) 104

Figure 4.10: Back vowels 106

Figure 4.11: Box plot of F1 and F2 by gender for English vowel [ɑː] in *pot* 107

Figure 4.12: Box plot of F1 and F2 by gender for English vowel [ɔː] in port 108

Figure 4.13: Box plot of F1 and F2 by gender for English vowel [ʊ] in *foot* 108

Figure 4.14: Box plot of F1 and F2 by gender for English vowel [uː] in food 109

Figure 4.15: Comparison of English vowel pair in *bit-beat* 111

Figure 4.16: Comparison of English vowel pair in *bet-bat* 111

Figure 4.17: Comparison of the English vowel pair in *bud-pot* 113

Figure 4.18: Comparison of English vowels pair in *pot-port* 114

Figure 4.19: Comparison of English vowel pair in *foot-food* 115

Figure 4.20: English vowel duration comparison based on gender 118

Figure 5.1: Vowel plot of Acehnese monophthongs 122

Figure 5.2: Vowel plot of Indonesian monophthongs 125

Figure 5.3: Plot of English [iː], Indonesian [i] and Acehnese [i] 128

Figure 5.4: Plot of English [ɔː], Acehnese [ɔ], and Indonesian [ɔ] 130

Figure 5.5: Plot of English [uː], Indonesian [u] and Acehnese [u] 131

Figure 5.6: Plot of English [ɛ], Indonesian [ɛ] and Acehnese [ɛ] 133

Figure 5.7: Plot of English [ʌ], Acehnese [ʌ], Indonesian [a] and Acehnese [a] 135

Figure 5.8: Plot of English [ʌ], Indonesian [u] and Acehnese [u] 136

Figure 5.9: Plot of English [ɪ], Indonesian [i] and Acehnese [i] 138

Figure 5.10: Plot of English [iː], English [ɪ], Acehnese [i] and Indonesian [i] 139

Figure 5.11: Plot of English [æ] Acehnese and Indonesia [ɛ] and Acehnese and Indonesian [a] 140

Figure 5.12: Plot of English [ɑː], Indonesian [ɔ] and Acehnese [ɔ] 142

Figure 5.13: Plot of English [ʊ], Indonesian [u] and Acehnese [u] 144

Figure 5.14: Plot of English [uː], Indonesian [u] and Acehnese [u] 145

Figure 6.1: Mean percentage of accuracy of the AX task 151

Figure 6.2: Mean percentage of accuracy based on NNB, SNA, SNB, and INA categories 153

Figure 6.3: AX Mean percentage of accuracy based on trial type 154

Figure 6.4: AX perception accuracy based on vocab size categories 155

Figure 6.5: Comparison of AX perception accuracy and lexical knowledge 156

Figure 6.6: Mean percentage of accuracy of ABX task 157

Figure 6.7: Mean percentage of accuracy based on NNB, SNA, SNB, and INA categories for ABX 158

Figure 6.8: ABX Mean percentage of accuracy based on trial type. 159

Figure 6.9: ABX perception accuracy based on vocab size categories 160

Figure 6.10: Comparison of ABX perception accuracy and lexical knowledge 161

Figure 6.11: Mean percentage of accuracy of FCI task 162

Figure 6.12: Mean percentage of accuracy based on NNB, SNA, SNB, and INA categories for FCI 163

Figure 6.13: FCI Mean percentage of accuracy based on trial type 164

Figure 6.14: Comparison of FCI perception score and vocab size categories 165

Figure 6.15: Comparison of FCI perception accuracy and lexical knowledge 166

Figure 6.16: Comparison of all perception task 167

Figure 6.17: Production of English vowels by an American speaker 171

Figure 6.18: Comparison of the production position of American and Indonesian speaker 173

Figure 6.19: Comparison of English [ɪ] - [iː] with equivalent Acehnese and Indonesian vowels 178

Figure 6.20: Comparison of English [ʊ] - [uː] with equivalent Acehnese and Indonesian vowels 180

Figure 6.21: Comparison of English [e] - [æ] with equivalent Acehnese and Indonesian vowels 181

Figure 6.22: Comparison of English [ɑː] - [ɔː] with equivalent Acehnese and Indonesian vowels 183

Figure 6.23: Comparison of English [ʌ] - [ɔː] with equivalent Acehnese and Indonesian vowels 184

LIST OF TABLES

Table 2.1: Acehnese Oral Monophthongs 44

Table 2.2: Acehnese Nassal Monophthongs 45

Table 2.3: Indonesian Monophthongs 46

Table 2.4: English Monophthongs (Ladefoged & Johnson, 2014) 47

Table 3.1: Word list for English monophthongs 72

Table 3.2: Word list for Acehnese monophthongs 73

Table 3.3: Vowel pairs and words for perception task 75

Table 3.4: ABX design for practice materials 77

Table 3.5: ABX design for actual materials 77

Table 3.6: AX design for practice materials 79

Table 3.7: AX design for actual materials 79

Table 3.8: FCI design for practice materials 81

Table 3.9: FCI design for actual materials 82

Table 3.10: Classification of speakers into high, mid and low group based on vocabulary size test 90

Table 3.11: Detailed distribution of vocabulary size score into high, mid and low group 91

Table 3.12: Word familiarity index for the high, medium and low group 92

Table 3.13: Detailed distribution of word familiarity index score for high, mid and low group 92

Table 3.14: Percentage and number of speakers familiar with words used in the study 93

Table 3.15: Word familiarity index by word pairs 94

Table 4.1: Average first and second formant measurements for English monophthongs and ED 96

Table 4.2: Classification of English vowel [ʌ] variations produced by Ach-IndE speakers 105

Table 4.3: Classification of English vowel /ʌ/ variation 110

Table 4.4: Duration of English monophthongs in a millisecond 116

Table 5.1: Average first and second formant measurements for Acehnese monophthongs 121

Table 5.2: Average first and second formant measurements for Indonesian monophthongs 124

Table 5.3: Classification of English vowels to Acehnese and Indonesian vowels 126

Table 5.4: Word familiarity index 136

Table 6.1: Vowel chart comparison across Acehnese (Asyik, 1987), Indonesian (Zanten, 1986) and American English (Ladefoged & Johnson, 2014) 148

Table 6.2: Comparison of English vowel pairs (VP) to Acehnese and Indonesian based on the framework by Flege (1987) 149

Table 6.3: Classification of English vowel pairs by NNB, SNA, SNB and INA categories 150

Table 6.4: Mean score and Standard Deviation for each English Vowel Pairs 152

Table 6.5: Mean score and Standard Deviation for ABX task 157

Table 6.6: Mean score and Standard Deviation for each English Vowel Pairs 162

Table 6.7: FCI trial type orders for [ɪ] – [iː] 164

Table 6.8: Average perception accuracy for the three tasks 168

Table 6.9: Formant values of English vowels by a male American speaker 170

Table 6.10: Formant values of English vowels by Acehnese-Indonesian speaker 172

Table 6.11: Prediction of English vowel contrast difficulty based on the American audio input 174

Table 6.12: Mean average of three perception task, AX, ABX and FCI 174

Table 6.13: Classification of English vowel pairs based on PAM in percentage 187

Table 6.14: Classification of English vowel pairs based on L2LP in percentage..... 191

LIST OF APPENDICES

Appendix A 241

Appendix B 242

Appendix C 247

Appendix D 252

Appendix E 257

Appendix F 258

Appendix G 259

Appendix H 260

Appendix I 263

Appendix J 264

Appendix K 274

Appendix L 284

Appendix M 291

Appendix N 292

Appendix O 299

Appendix P 306

Appendix Q 312

Appendix R 314

Appendix S 315

LIST OF ABBREVIATION

MIBS	: Modern Islamic Boarding School
Ach-IndoE	: Acehnese Indonesian English
MalE	: Malaysian English
BruneE	: Brunei English
SingE	: Singapore English
ThaiE	: Thailand English
VP	: Vowel Pair
ED	: Euclidean Distance
SD	: Standard Deviation
FCI	: Forced Choice Identification

CHAPTER 1: INTRODUCTION

1.1 Introduction

This chapter provides the background to the study leading to the research problem and gap which has motivated this study. Following this, the objectives and research questions are presented. The significance and scope of the study are subsequently presented, followed by the outline of this thesis.

1.2 Languages in Indonesia

Indonesia is a multilingual country where most citizens speak more than one language, a vernacular language and Bahasa Indonesia (BI) (Sobarna, 2007). Bahasa Indonesia is the national language and is used in formal settings such as in government agencies, businesses, and the media (Alwi & Sugono, 2000; Perwitasari, 2018). Bahasa Indonesia is also the medium of instruction in schools and universities, and the main language of communication in formal letters, legal documents, the media, and government or business documents. Both government and private employees use Indonesian as the main medium of communication in their organisations.

Bahasa Indonesia belongs to the Austronesian family of languages. The language is of Malay origin, the old trading language or lingua franca used along the Straits of Malacca that is also spoken in Malaysia, Singapore, and Brunei (Kirkpatrick, 2010). Thus, Bahasa Indonesia share the same root to Malay spoken in Malaysia, Singapore, and Brunei During the sixteenth century when Islamic kingdom ruled part of Indonesia, Malay emerged as the most important language of trade and Islam (Arsip Nasional Republik Indonesia, 2023). Before Indonesia gained independence, the language was established as an identity marker to unite Indonesians through the Indonesian Youth Declaration in 1928 (Raditya, 2021). In this declaration, the youth from all over Indonesia gathered in

Jakarta and pledged to use Bahasa Indonesia as the unity language. By the time Indonesian gained Independence in 1945, Indonesian was declared as the national language and gradually gained popularity as the first language among Indonesians living in the city. Ir. Soekarno, the first Indonesian president, also read the declaration of Indonesian Independence in Bahasa Indonesia to signify the importance of the language for the country. After Independence, it was reported that only around 3% of Indonesians spoke the language, but the number of Indonesians speaking Bahasa Indonesia significantly increased to 34% by 2000 (Montolalu & Suryadinata, 2007, p. 47). In another report, it was argued that the figure was much higher after Independence. Errington (1992, p. 419) reported that the number of Indonesians speaking Bahasa Indonesia had reached 60% by 1992, while Gordon (2005) reported that the figure had increased to 68% (140 million) by the year 2000. Quinn (2001) predicted that by the 21st century, the number of Indonesians learning or speaking Bahasa Indonesia as their first language would reach 80% of the total population.

In addition to Bahasa Indonesia, local languages are used informally such as in families, at local markets, at traditional ceremonies, and among neighbours. It is recorded that at least 700 local languages are actively spoken in Indonesia (Mabruri, 2021) spreading across its 37 provinces (Kusnandar, 2022) and the five main islands of Sumatra, Java, Borneo, Celebes and Papua. The languages are used by an estimated 1300 ethnic groups (Yuniarni, 2016). Approximately 10 vernacular languages are widely spoken: Javanese, Sundanese, Madurese, Minangkabau, Musi, Buginese, Banjarese, Acehnese, Balinese and Betawi (Khan, 2020). Javanese, Sundanese and Madurese are spoken by 84.3 million, 34 million, and 13.6 million speakers respectively (Khan, 2020). These three languages are mainly used on the island of Java. The rest of the languages have less than six million speakers. Acehnese is the sixth most spoken language on par with Buginese and

Banjarese, with 3.5 million speakers, and it is positioned as the second most spoken language outside Java (Khan, 2020).

Having acquired a vernacular language and/or Bahasa Indonesia, for most Indonesians, English is not a second language but is likely to be the third language learnt as a subject at the secondary school level. English is the only compulsory foreign language subject taught from the ages of 10 to 17 years old, that is from the first year of secondary school until the third year of senior high school (Hadisantosa, 2010; Kemendikbud, 2014; Kirkpatrick & Sussex, 2012; Mistar, 2005). Even though English was once taught at the primary school level, starting from Grade 4 to Grade 6 (Renandaya, 2000), Indonesia is now the only country in Southeast Asia that does not make English a compulsory subject from the primary school level (Kirkpatrick & Sussex, 2012). Indonesians continue to learn English in higher education as a compulsory general subject, which is worth at least two credit hours. Some universities also require students to pass English proficiency tests as a graduation requirement.

In this regard, the status of English as a foreign language in Indonesia contrasts with neighbouring countries such as Malaysia, Brunei, and Singapore. In Malaysia, despite Malay being the national and official language and the main medium of instruction in public school education (Pillai et al., 2021), "English is given recognition in the education policy and planning as the 'second language'" (Thirusanku & Yunus, 2014, p. 225). This status of English is different from Brunei and Singapore despite the three countries being former British colonies. Constitutionally, Brunei designated Malay as the national language and English as the official language (Sharbawi, 2021). In Brunei, Malay is only used as the medium of instruction for three years in primary schools, after which the students learn mathematics, science, geography, history, and technical subjects in English while the rest are taught in Malay (Kirkpatrick, 2012). In contrast, Singapore recognises

four official languages English, Malay, Mandarin, and Tamil, with Malay being the national language (Tan, 2007). However, in reality, English has become the most dominant language in Singapore, with the government introducing an "English + 1" policy in which Singaporeans must learn English and a local language to preserve their local identity (Kirkpatrick, 2012, p. 31). Thus, English is widely used in Singapore and is spoken to a large extent in Malaysia, and in both countries, a local variety of English has developed from the time it was transplanted to these countries. In contrast, Bahasa Indonesia does not have a large English speech community, and its use is primarily confined to educational settings and international communication. In other words, English remains a foreign language in Indonesia.

Given the multilingual nature of Indonesia, studying the production and perception of English among Indonesians should take into consideration the fact that most Indonesians may already speak more than one language before learning English. The languages in which they are bilingual or multilingual may differ based on where they live and grow up in Indonesia. It would, thus, be naïve to view Indonesians as a homogenous group as if Javanese/Indonesian and Acehnese/Indonesian bilinguals were somewhat identical and referred to generically as Indonesian bilinguals. Even though most of the vernacular languages in Indonesia belong to the Austronesian language family, their vowel systems, for example, are likely to differ considerably. For example, Javanese, which is classified as Malayo-Javanese, only has six oral vowels (Perwitasari, 2019), while Acehnese, which is a member of the Malayo-Chamic language family, has ten oral vowels. Acehnese not only has a bigger vowel system than Javanese but has more vowels that are similar to English vowels. In comparison, Bahasa Indonesia only has six monophthongs (Soderberg & Olson, 2008).

Due to larger vowel inventories compared to Javanese and Sundanese, Acehnese/Indonesian bilinguals may perceive and produce English vowels differently from Javanese and Sundanese bilinguals (Perwitasari (2019), and might be able to benefit from the larger vowel inventory in Acehnese which has more vowels similar to English than Bahasa Indonesia. Previous studies have reported that learners with larger native phonemic inventories tend to learn non-native vowels easier (Iverson & Evans, 2007). Studies have also indicated that speakers who have more native vowels, that is vowels in their first language, similar to the target language vowels are able to refer to their existing vowels when learning a new language (Alispahic et al., 2017). In view of these studies, it is posited that Acehnese-Indonesian bilingual speakers may make use of their existing vowel inventories in both these languages when producing English vowels.

1.3 English in Islamic Boarding Schools

As stated in the previous section, English is not widely spoken outside educational and international settings. In an educational setting, the use of English tends to be limited to the English language classroom context. In addition, many teachers use Bahasa Indonesia in the teaching of English (Zein, 2017). This may explain why the number of people who are fluent in English remains low in Indonesia (Kirkpatrick, 2010). Thus, the use of English in daily communication is rare in Indonesia. However, at present, some schools in Indonesia promote the use of English outside the classroom context. These include Modern Islamic Boarding (MIB) schools, which are fully residential schools comprising male and female students aged 16 to 18 years old.

MIB schools have a unique policy requiring students to speak English after school hours within the boarding school area, such as at the school dormitory, sports field, mosques, and canteen. Failing to comply with these requirements results in the students being punished, which may entail carrying out community services or even being expelled from

the school. Such a requirement forces students to speak English despite their lack of proficiency in English to avoid punishment.

The constant use of English within the confines of the boarding school during their three-year stay at the school may lead to students developing particular features of English unique to their boarding school. Due to limited resources and authentic English input, the students mostly rely on their dormitory teachers often called *ustadz* 'teachers', seniors and peers for input. *Ustadz* are mostly alumni from similar MIB schools and have an adequate command of English proficiency to speak with the students. They would have had some experience communicating in English when they were students. *Ustadz* may come from different backgrounds of study. However, the MIB school often hires at least one male and female dormitory teacher who are students of or have graduated with education from an English department. These *ustadz* become the benchmark for the students to rely on as authentic input. With such limited authentic input and being forced to speak English for everyday communication despite the natural tendency to use Acehnese or Bahasa Indonesia in such contexts, their production of English may be influenced by the languages they speak, which are Acehnese and Bahasa Indonesia. This may be apparent in their pronunciation and other linguistic features inherent in the English they produce. This is to be expected because when someone learns a second language (L2), the first language (L1) is said to act as the first state of the learners' mental structure on which the subsequent language to be filtered through (Hummel, 2013), while when they acquire a third language (L3), such as English, they would already have a combination of L1 and L2 phonemes in their vowel inventories (Amaro, 2012). Since their immediate resource of phonemes, such as vowels, when producing English are Acehnese and Bahasa Indonesia, they may transfer their knowledge of existing L1 and L2 vowels into English. In relation to this, Escudero (2005) highlights that when learning a second or third language, learners usually filter non-native vowels through their existing

vowel system. As a result, the production of English vowels produced by the speakers from this speech community may be produced with reference to similar Acehnese and Bahasa Indonesia vowels.

1.4 Problem Statement

This research was initially motivated by the researcher's past experiences and observations as a student and teacher at one of the Modern Islamic Boarding (MIB) Schools in Banda Aceh. It was observed that the students at the MIB school could communicate in English in their daily conversations in the school, and their pronunciation is, as might be expected, distinct. It seemed as if their production of English vowels were similar to equivalent ones in Bahasa Indonesia and Acehnese. This initial observation led to the assumption that Acehnese-Indonesian (henceforth, Ach-IndE) speakers at the school may have developed particular features of English vowels which may represent an Acehnese-Indonesian variety of English. This variety may be different from the English produced by speakers in other parts of Indonesia from a non-MIB school context. One of these features is their production of English vowels.

Such potential variety is of relevant to be viewed under the World English paradigm. In World Englishes paradigm, Indonesia is placed in Kachru's Expanding Circle (Bolton et al., 2020; Kachru, 1998; Kirkpatrick, 2010), which is a norm-dependent country. As a norm-dependent country, Indonesia has yet to develop a unique variety of English and is open to adopting established varieties from the Inner Circle (such as British and American English) or the Expanding Circle (such as Malaysian English and Singapore English) (Kachru, 1992a, 1992b). Due to the function and status of English as a foreign language in Indonesia, Bolton et al. (2020) do not use the term Indonesian English but prefer to refer to it as English in Indonesia. This term denotes that Indonesia has not established a unique variety distinct enough to be called Indonesian English. However, if ones look at

the English practised by students at MIB schools, where the students are required to use English in their daily communication, a particular sub-variety may well be developing in this context. Since MIB schools are scattered across provinces in Indonesia, their features may be influenced by the different vernacular languages that the students speak. Since Indonesian students also speak Bahasa Indonesia, some pronunciation features may also be similar across all Indonesian speakers. First-year students at MIB schools have limited or no exposure to English conversation prior to enrolling at these schools. This is because, as previously mentioned, English is not a compulsory subject in primary schools. As a result, when the schools require them to speak English after three months of being at the school, these students tend to rely on their first language when speaking. This is mostly apparent in their choice of words and pronunciation.

Most studies have been directed toward comparing the production of L3 phonemes to native speakers in order to measure their proficiency (Larson-Hall, 2006; Perwitasari, 2019; Riney & Takagi, 1999; Wrembel, 2010). Such direction has ignored the fact that English is a lingua franca that is used not only to communicate with native speakers of English but also with speakers of various L1 as in the international university setting. The uniqueness of regional accents has been given less attention. Few studies have positioned the characteristics of English phonemes produced by L3 learners of English as distinct and unique without any comparison to "standard" phonological sounds. Instead of comparing regional English varieties to "standard" English (Deterding & Kirkpatrick, 2006; Perwitasari et al., 2015), it would be prudent to describe their English production on their own to avoid treating them as deviant from a standard variety (Pillai, 2014, pp. 58-59).

In the context of Indonesia, there have been several studies on the production (Fata et al., 2017; Low, 2016; Perwitasari et al., 2016; Perwitasari et al., 2015) and perception of

English vowels (Perwitasari, 2018) by Indonesian speakers. However, none of these studies compared the production of English vowels to those in Indonesian or other regional languages to explore the degree of similarity among the vowels produced. Thus, the influence of existing Bahasa Indonesia and regional vowels on the production of English vowels remains unexplored.

In another study Low (2016) only recorded the data produced by one Indonesian speaker who has lived in Singapore for quite a while. This makes it difficult to generalise that the Bahasa Indonesia vowels produced represent how all Indonesians produce English vowels, considering Indonesians with different vernacular languages may produce English vowels differently. Even though the studies by Perwitasari (2019) and Perwitasari et al. (2016) explored the production of English vowels by Javanese and Sundanese speakers, no analysis was made on the possible effect of Javanese and Sundanese on English vowels. Furthermore, Fata et al. (2017) investigated the production of English vowels by Acehnese high school students, but the participants only used English during English subjects, and no data on proficiency was provided. The quality of their English vowels may be attributed to a lack of proficiency due to the limited usage of English in conversation.

Thus, previous studies on the production of English vowels by Indonesian speakers did not investigate the effect of the vowels in Bahasa Indonesia and/or regional languages on the production of English vowels among Indonesian speakers. This is the gap that this study aims to begin filling; in fact, as mentioned in 1.2, it has been argued that learners tend to filter non-native vowels through their existing L1 and L2 vowels (Best & Tyler, 2007; Escudero, 2005; Flege, 1995) which might have an effect towards the production of non-native vowels, which may well be the case of Acehnese-Indonesian speakers at MIB schools.

In terms of perception, thus far, only two studies have looked at the perception of English vowels (Perwitasari, 2013, 2018, 2019), but these studies did not make any comparison to their production of the same vowels in English nor did it attempt to determine the perception accuracy of English vowel contrast. Instead, it focused on the error rates of English vowel perception. So, to what extent the English vowel contrast perceived and produced by Indonesian speakers are related was left unexplored. The study used mouse tracking to determine the error rates of the perception of individual English vowels instead of the vowel contrasts. As a result, the perception accuracy of English vowel contrast by speakers was not provided. Thus, the current study also seeks to determine the perception accuracy of English vowel contrasts by Acehnese Indonesian speakers and how the accuracy corresponds to the production of English vowel contrasts.

1.5 Research Aim and Objectives

In view of this research gap, this study aims to look at the English monophthongs produced and perceived by Acehnese speakers of English, who have considerably more English language input because of their school's policy of English language use. This study focuses on the production and perception of English monophthongs and does not include diphthongs and consonants because English monophthongs is the pure form of vowel and while diphthongs vary greatly across English varieties. The objectives of this study are as follows:

1. To examine how the Acehnese-Indonesian speakers in a Modern Islamic Boarding school in Aceh produce English vowels.
2. To compare the English vowels with Bahasa Indonesia and Acehnese vowels produced by the speakers.
3. To determine how the speakers perceive English vowels.

4. To examine the relationship between the perception and production of English vowels produced by the speakers.

1.6 Research Questions

Based on the research objectives, the current study seeks to answer the following research questions:

1. What is the quality of the English monophthongs produced by the Acehnese-Indonesian speakers at the selected Modern Islamic Boarding school in Aceh?
2. To what extent do Acehnese-Indonesian speakers distinguish typical English vowel pairs in terms of vowel quality and length contrast?
3. To what extent are equivalent vowels in English, Bahasa Indonesia, and Acehnese produced similarly by the speakers?
4. How do Acehnese-Indonesian speakers perceive English monophthongs?
5. To what extent is their production of the English vowels related to their perception of the same vowels?

1.7 Significance of Study

While, there are a number of previous studies looking at vowel production (Perwitasari et al., 2016; Perwitasari et al., 2015) and perception (Perwitasari, 2018) of English vowels by Indonesian speakers, studies on Acehnese-Indonesia speakers is scarce (Fata et al., 2017). Thus, this study intends to start filling this research gap and contribute to our understanding of how bilingual speakers produce vowels in a third language. It also aims to contribute to our understanding of the perception and production of English vowels by Acehnese-Indonesian speakers.

In addition, most of the previous studies compared the production of English vowels to American speakers (Perwitasari et al., 2016; Perwitasari et al., 2015) as a target instead of treating it as a unique variety. Thus, this study has the potential to contribute to the discussion about the features of English in Indonesia within the framework of World Englishes.

1.8 Scope of Study

This study focuses on the production and perception of English monophthongs and does not include diphthongs and consonants. Only the production of monophthongs across three languages, Acehnese, Bahasa Indonesia, and English, is covered in the current study. The participants of the study are confined to one Modern Islamic Boarding school, who can speak Bahasa Indonesia and Acehnese and originate from West Aceh. This study does not cover all ethnic groups in Aceh and does not aim to generalise the findings to other Islamic boarding schools in Aceh. Different levels of English proficiency from different English speech communities in Aceh may have different English features, which this study does not account for.

1.9 Thesis Outline

This thesis is presented in seven chapters. Chapter 1 introduces the study by providing an overview of the study: its background, problem statement, research objectives, research questions, as well as the significance and scope of the study. Chapter 2 provides the literature review and the conceptual framework of the study, while Chapter 3 describes the methodology of the study. Chapter 4 presents the results for the production of English monophthongs, while Chapter 5 contains the findings for the comparison of English monophthongs to equivalent Acehnese and Bahasa Indonesia monophthongs. Chapter 6 presents the findings for the perception of vowel contrast in English and the relationship

between the perception and production of English vowel contrasts. Chapter 7 provides a discussion of the findings from Chapters 4 to 6, and Chapter 8 concludes this thesis.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter begins by discussing the World Englishes paradigm and then moves on to present the position of English in Indonesia in comparison to neighbouring countries. A review of studies on the production and perception of English vowels in the Indonesian context and the neighbouring variety of English is presented.

2.2 Global Englishes as a Theoretical Framework

Many scholars have attempted to describe the spread, status, and function of English in various countries and contexts around the world (Jenkins, 2014; Kachru, 1992a; Schneider, 2007). The status and function of English worldwide can no longer be explained simply using the traditional models where the term 'native' speakers and 'standard' English belittle the English speakers growing up with English as their first language and the wide varieties of English spoken outside the United Kingdom, the United States, Canada, Australia and New Zealand. Two major models have been proposed in regard to World English, Schneider's Dynamic Model (Schneider, 2007) and further work on Extra- and Intra-territorial Forces (EIF model) (Buschfeld & Kautzsch, 2017), Kachru's Three Circles (Kachru, 1992a, 1998) and one influential model under Global English, Jenkin's English as a Lingua Franca (Jenkins, 2014). Sadeghpour and D'Angelo (2022) suggested that Global English should be used as the umbrella term for World English and English as a Lingua Franca Paradigm. Thus, this study adopted the Global English term as the umbrella for the three models to be discussed. This section first discusses Kahcru's (1988) Three Circles model under the World Englishes paradigm as the framework for the present study. Two other models, Schneider's Dynamic Model and the concept of English as a Lingua Franca are also evaluated.

2.2.1 Kachru's Three Circles Model

The World English paradigm was proposed by Kachru in an effort to promote the many English varieties (India, Singapore, Nigeria) around the world outside the traditionally known English-speaking countries (the United Kingdom, the United States of America, Australia, Canada, New Zealand). With the spread of English worldwide and the use of English as a global language, Kachru (1992a) called for a shift in paradigm to recognize pluralism in English. Kachru argued that when teaching English, English teachers should consider whether students would use English for international comprehension (such as in Japan) or intranational communication (such as in India) (Kachru, 1992b). He challenged the traditional ways of dividing the status of English in certain countries as English as a Native Language (ENL), English as a Second Language (ESL), and English as a Foreign Language (EFL) and proposed the popular Three Circle model to recognize the diverse varieties of English used worldwide and the function of English served in various communities.

The Three Circles model challenged the status quo in the use of English as a native language for countries such as the United Kingdom, the United States, Canada, and New Zealand and the demeaning use of English as a Second Language for countries such as India, Singapore, Malaysia, and Nigeria. English was never considered native in these countries despite the fact that many current generations in Singapore speak English as their home language (Kirkpatrick, 2010). The term "native" used for speakers whose first language is English from the United Kingdom, the United States, and Canada, is deemed no longer applicable, considering people in other parts of the world were born with English as their first language. In Singapore, for example, 60% of primary school children spoke English as their home language by 2009, compared to only 40% speaking Mandarin as their home language (Kirkpatrick, 2010). Using the traditional way of dividing countries around the world, Singapore would be placed under countries with English as a

Second Language. This classification would discriminate against the identity of Singaporean children who grew up with English as their first language. In relation to this, Schneider (2007) argued that native speakers of English could be found not only in Canada and New Zealand but also in Singapore and Nigeria. Thus, in place of the traditional ways of dividing English around the world, (Kachru, 1988, 1992a) proposed that countries around the world should be divided into three categories called "circle" Inner Circle, Outer Circle (Extended), and Extending Circle. Kachru (1988, p. 5) argued that "these circles represent the types of spread, the patterns of acquisition, the range of functional domains, and the societal penetration of the language."

First, the Inner Circle refers to countries where most citizens speak English and is used in almost all aspects of conversation in the country. The countries of the inner circles are norm-providing nations that provide norms (English standard) to other countries, especially Expanding Circle. This circle is dominated by countries in which English is the mother tongue. These include the USA, Canada, the UK, Australia, and New Zealand. However, he did include South Africa and Jamaica because the sociolinguistic context of English status in these countries is complex to the rest of the Inner Circle (Kachru, 1992b). The complex situation that Kachru refers to here is due to the status of Jamaican English as an English creole, similar to that in Nigeria (Schneider, 2007). Schneider (2003) also challenged that placing the United States, Australia, and New Zealand in the inner circle seems to overlook the minorities in these countries, such as Native Americans and Hispanics in the United States of America, Aboriginal communities in Australia, and the Maoris in New Zealand.

Second, Outer Circle countries are where English has been institutionalized as an official or additional language. These countries include Singapore, Malaysia, Brunei, the Philippines, Sri Lanka, and India, to name a few. English is an institutionalized language

in this country in addition to the local or national language (Kachru, 2005). For example, Singapore has Malay as the national language and four three official languages (English, Chinese, and Tamil) (Kirkpatrick, 2010). Outer Circle countries are norm developing and have developed their own unique variety of English, such as Singapore English and Indian English (Kachru, 1996). Singaporean English has been comprehensively documented by Deterding (2007), and in fact, English has gained traction to become a dominant first language in Singapore (Kachru, 2005). A 2001 survey reported that 40% of Singaporean children speak English as their home language (Tan, 2007), and the figures had risen to 60% by 2009 (Kirkpatrick, 2010). Such a high usage of English as the home language among Singaporean children begs the question of whether Singapore should still be considered an Outer Circle variety.

Finally, Expanding Circles are countries where English is the primary foreign language taught in schools. These include all countries outside the Inner and Outer Circles where English has gained popularity (Kachru, 1992b). Most people in these countries are English-knowing individuals who learn English at school. Due to the absence of speech communities in the Inner Circle and Outer Circle countries, speakers in the Expanding Countries may never use the language outside the classroom context. In the context of South East Asia, Indonesia, Thailand, the Philippines, Vietnam, and Cambodia fall under this category (Kirkpatrick, 2012). Countries in expanding circles usually have a common national language that people use to speak with each other and do not institutionalize English as an official language. Expanding circles are similar to what was referred to as English as a Foreign Language (EFL) (Schneider, 2007). In expanding circles, English is learned as a foreign language; for some, it is merely a subject in schools or universities. Since English has gained popularity in these countries and is considered a must-have skill by many, some schools even promote the use of English as the language of instruction as in the case of bilingual school (Hadisantosa, 2010). Due to these conditions, these

countries are said to be norm-dependent which means they are ready to accept the English norms from the countries in the Inner Circle.

However, even though many scholars have widely adopted the model (including Jenkins (2013) and Schneider (2014) when describing the spread of English, much of Kachru's work in World Englishes heavily focused only on Inner Circle versus Outer Circle varieties. Many references have been made to new emergent varieties of English in Outer Circles (Deterding, 2007; Mohamad & Deterding, 2016; Pillai et al., 2010), but little explanation is given to countries in the Expanding Circle as emergent varieties. Kachru placed countries such as Indonesia in expanding circles and rejected the term English as a foreign language for these countries (Kachru, 1992b) but he has not provided a proper means of describing Englishes outside outer circles. As a result, scholars were in disagreement when referring to English varieties in the Expanding Circle. Some treated them as similar to Outer Circle varieties as emergent varieties such as Japanese English (Lee & Low, 2021), Chinese English (Sundkvist & Gao, 2016), and Indonesian English (Aziz, 2003) while others preferred a conservative term such as English in Vietnam (Sundkvist & Nguyen, 2020), English in China (Deterding, 2006) and English in Indonesia (Lauder, 2020). Scholars using the World Englishes paradigm followed through by referring to Englishes in the External Circle not as emergent varieties but as English spoken by the people within a specific country. Bolton et al. (2020) in the Handbook of Asian Englishes also never used the term Country X English but English in Country X to refer to Englishes spoken outside the Inner and Outer Circle countries.

These conflicting references to Englishes spoken in the Outer Circle were heavily indebted to how Kachru (1992a) provided the characters of Englishes in each circle. In the Three Circles Model, Inner Circle countries are norm-providing countries, providing English norms to other countries, including to the Outer Circle countries. Outer Circle

countries are norm-developing countries that are developing their own varieties, unique and distinct from the norm-providing countries. Expanding Circle countries are norm-dependent and have yet to establish their own variety and are open to being influenced by the varieties from Inner Circles and Outer Circles. However, in a more recent publication, Kachru and Nelson (2006) acknowledged that Outer and Expanding Circles shared some characteristics of Englishes while Low (2016) found that Expanding Circle countries exhibited similarities toward both Inner and Outer Circles. Thus, Davydova (2012, p. 383) argued for an integrated approach across both Circles, proposing that "robust generalizations can be made across learner Englishes and indigenized L2 Englishes."

A more robust and integrated approach to English in Expanding Circles was recently proposed by Buschfeld and Kautzsch (2017), in the *Extra- and Intra-territorial Forces Model*. This model does not necessitate the colonial presence in determining the status of English in a country. The colonial presence restricts the Dynamic (Schneider?) and Kachru's Three Circles' model when it comes to capturing English used in countries without colonial past (Bruthiaux, 2003). In place of a colonial presence, the model by Buschfeld and Kautzsch (2017) argue for globalisation as the initial force of English usage in countries without colonial presence such as those in Europe. For example, urban youth in Finland have fully embraced English and English has a significant role in their life (Leppänen et al., 2011). Similar behaviours were found among German youth who are reluctant to embrace their national identity due to German's historical legacy and instead use English to signify European and global identity (Erling, 2004). In addition, the model also questions the ESL-EFL dichotomy when both varieties of English were found to share similar features. For example, Nesselhauf (2009) compared ESL and EFL varieties and found that new prepositional verbs (e.g. discuss about, enter into, request for) that were often regarded as innovation in ESL variety has been considered 'errors' in

EFL variety. Such differences in referring to the same shared features of ESL and EFL contributed to the dichotomy of both varieties in World English paradigm.

In the case of Indonesia, it is unclear if the variety of English in Indonesia (if one exists) should be considered a new variety and labelled as Indonesian English or English in Indonesia. Considering that English is a foreign language in Indonesia and is mainly taught as a subject in schools, a speech community that gradually converses in English is rare. However, as mentioned in the previous chapter (see section 1.3), some schools in Indonesia promote English as the language of instruction and communication, such as the Modern Islamic Boarding School (MIB). Thus, even though the speech community is confined to a particular school community, English spoken by these speakers may exhibit a general pattern due to the influence from Bahasa Indonesia and Acehnese and peer influence. Kachru (1992b) termed this pattern 'bilingual's creativity. In bilingual creativity, the difference was not seen as inadequacy in English but as characteristics and features of a particular variety of English. In this view, the speech was adjusted to accommodate the psychological, sociological, and attitudinal upbringing of the culture where the language is in contact (Kachru, 1985), which is in this context with Bahasa Indonesia and Acehnese. The constant use of English as the language of communication in MIB schools may encourage learners to exhibit a specific pattern of pronunciation. MIB schools, in particular, rarely have native speaker teachers as the model. The students rely mostly on teachers, seniors, and friends who are Indonesian. As a result, their English pattern (especially the pronunciation and accent) may differ from English spoken in Inner and Outer Circles (see section 1.3 for further context on English policy in MIB schools).

For the current study, the Acehnese-Indonesian speakers do not learn English in an L2 environment, and their use of English is not confined to the classroom context. They learn English to communicate with their peers, seniors, and teachers during their stay at school

and use the language regularly except during classroom hours where Bahasa Indonesia as the medium of instruction. Due to the lack of native speakers' input because their role models are seniors and teachers (who are locals), their perception and production of English are heavily constrained by their perception of L1 sounds. This condition results in their English developing a unique English variety which may be different from those spoken in Inner Circle and Outer Circle countries. Since World English is aimed to account for the many unique varieties of English, the paradigm fits the phenomenon of English spoken by students at Islamic boarding schools. The World Englishes paradigm does not prescribe the English from Inner circle countries as standard as they do in theories of Second Language Acquisition. World English treats an English variety as a possible emergent variety and comparison is mostly made to other varieties from the same country or established emergent varieties from adjacent countries such as Sharbawi (2012), Pillai (2014), and Low (2016).

2.2.2 Schneider's Dynamic English Model

Schneider's Dynamic Model (Schneider, 2003, 2007, 2014) tried to classify the status of English outside the United Kingdom based on its developmental stages. These countries were British colonies that had gained independence or retained sovereign status. The model was developed based on the theories of language contact, sociolinguistics, social identity, and language evolution (Schneider, 2014). He proposed that postcolonial English undergoes five stages to become a new variety (Schneider, 2003). The phases he proposed were *foundation, exonormative stabilization, nativization, endonormative stabilization,* and *differentiation.* This model is relevant to the current study because it also attempts to describe the development of new English varieties. However, the new varieties of English included under this model were limited to Outer Circle countries due to the requirement it has set which includes the foundation level. At the *foundation* phase (Phase 1), the model describes that English was introduced by migrants in the newly

colonized territories. Dialect contacts and *koinezation* took place at this stage because the settlers may have come from a diverse regional dialect from their original country. Pidginization and toponymic borrowings are also likely to occur, while some local people may be exposed to a limited amount of English. Countries in the Expanding Circle do not have a large number of migrants coming to these countries, which makes the *foundation* phase of the model inapplicable to the model. This view was later highlighted by Schneider (2014) who states that the Expanding Circle countries do not have a foundation stage and start from Phase 2, *exonormative stabilization*. Thus, based on the *foundation* characteristics, Indonesia, as an Expanding circle country did not go through the foundation stage.

However, even in Phase 2, English spoken in the Expanding Circle countries does not totally conform to the characteristics specified in the model. In Phase 2, the model states that *exonormative stabilization* develops when the contact between settlers and the indigenous community increases due to political stabilization. The norm at this stage was to accept the external features of British English. However, lexical borrowing (in term of fauna and flora) and transfer (phonology and syntax) from the second language users were increasingly discovered at this stage. For the Expanding Circle countries, Schneider (2014) argued that at this stage, the countries exhibit five indications, use in higher education, use in other formal context, exonormativity, widespread bilingualism and cultural borrowings. However, none of these indications can be fully observed in Indonesia. Even though some international private universities in Indonesia adopt English as their language of instruction, the practice is limited to urban cities, the number of such universities is scarce, and only a handful of elites go to these universities. Nevertheless, Ike (2012) still considers Japan to have gone through Phase 2 despite its lack of the features specified in the Dynamic Model.

The next three phases specified in the model are also not applicable to the context of English development in Indonesia. In Phase 3 *nativization*, English started to experience cultural and linguistic transformation. This happened because the social gap between the descendants of original settlers and local people had been reduced. Increased interaction among this generation results in the emergence of a newly developing variety (words, pronunciation, and grammar). Even though these habits were salient among the local descendants, the descendants of original settlers were starting to embrace the trend. Malaysia and Singapore are an example of countries that have reached Phase 3. The next stage in Schneider's model is *endonormative stabilization* (Phase 4). This stage was usually marked by political independence when the new generations started to develop a sense of belonging to their new country. Sometimes catastrophic events called Event x, also contributed to this stage where the descendants tried to break free from the mother country. An example of the country undergoing this Phase is South Africa and India. New variety at this stage was initially promoted with the promotion of local dictionaries and grammar. From here, the newly established variety went on to *the differentiation phase* (Phase 5). Australia, Canada and New Zealand are examples of countries that have reached this final phase in the Dynamic Model. Internal stabilization and cohesiveness led to the establishment of a new emergent dialect within the new variety. This resulted in different places within a country speaking new dialects locally.

Looking at the phases offered in the Dynamic model, countries included in this model all had gone through the *foundation* and *exonormative stabilization* phase. This prerequisite restricts countries with no colonial history with the British such as Indonesia, from undergoing the *foundation* and *exonormative stabilization* phase. Most countries where English was traditionally called a foreign language are excluded from the Dynamic Model. As a result, the development of English in Indonesia could not be explained based on the five phases offered by (Schneider, 2003, 2007). Even though the British colonized

Indonesia for a short period from 1811 to 186 (Nailufar, 2022), the settlement was mostly on Java island, and the *exonormative stabilization* stage never reached its maturity because Java was given back to the Dutch.

However, in 2014, (Schneider) extended the model to re-evaluate the possibility of the Dynamic Model covering countries in the Expanding Circles. Even though some scholars (Buschfeld, 2013; Ike, 2012) have attempted to explain English development in the Extending Circles through the lens of the Dynamic Model, Schneider (2014) argued that these countries did not undergo Phase 1 and have yet to fully exhibit some components in Phase 2. However, in Extra- and Intra-territorial Forces (EIF model) Buschfeld and Kautzsch (2017) argued that the foundation phase can be in the form of ESL in which colonization (direct or indirect through media) is viewed as the Phase 1. This model is the most comprehensive model that took at explaining the gap between post-colonial and non-post-colonial countries. The Phase 2 components include the use of English in higher education and other formal contexts, exonormativity, widespread bilingualism, and cultural borrowings. East Asian Expanding Circle countries, for example, show weak parallels to Phase 2 due to the small number of bilingualism, limited use of English in a formal context and any characteristics of later developmental stages are yet to be found (Schneider, 2014). Hence, at the end of the paper, he proposed that a different conceptualization should be used to account for the dynamic development of English beyond the Inner and Outer circle countries. A concept that Kachru predicted decades ago is that English should be discussed beyond the language's origins and 'native speaker' centeredness. Englishes in expanding circles are now viewed as an economic resource (Kachru, 2005) and could move to become a multicultural resource detached from the Western cultural context (Schneider, 2014), as in the case of promoting English to teach Islamic values in some Modern Islamic Boarding schools in Indonesia.

2.2.3 Jenkin's English as a Lingua Franca

The model of English as a Lingua Franca (ELF) was advocated by Jenkins (2000) from her experience in supervising postgraduate students from Asia studying in the United Kingdom and Seidlhofer (2004). According to Jenkins (2013), the term English as a Lingua Franca is not new to the academic world and has been used for many centuries in various forms. English as a Lingua Franca had been used interchangeably with international English and English as an International Language and was often misunderstood by World English (WE) scholars as advocating the standardized form of English (Jenkins, 2006). In resolving the confusion about what ELF means, Jenkins (2013, p. 24) defined the term as "English when it is used as a contact language between people from different first languages (including native English speakers)". In another publication, Deterding et al. (2013) also defined EFL based on the communication setting. In this book, she defines EFL as "English used by speakers from postcolonial Outer-Circle countries … as well as Expanding-Circle … when they are conversing with speakers from other countries in the Outer or Expanding Circles" (Deterding et al., 2013, p. 1). In addition, ELF also made a very insightful distinction from English as a Foreign Language (EFL). Seidlhofer (2013) argued that ELF differs from EFL in that the objective of EFL is to become integrated with Native speakers while EFL focuses on intelligibility and communication among nonnative speakers of English (NNE) and with native speakers of English (NE). EFL encourages adoption and imitation of native speakers, while ELF appreciates accommodation and adaptation. ELF also differs from WE in that WE scholars focused on "bounded varieties, that is the nativized Englishes of post-colonial nations" while ELF "operates across national boundaries" (Jenkins, 2013, p. 27).

Based on the definition, the context of EFL is associated with international encounters among speakers of various L1 communicating in English, which mostly happens in academic and business settings. In the case of the current study, the setting of

communication takes place in a Modern Islamic Boarding school among Indonesian students who share the same national language, Bahasa Indonesia. The present study setting is not international since the speakers are local students from Aceh. At these schools, English is selected as the means of communication as part of an education program to teach and promote English to students. Students and teachers are both Indonesian and speak Bahasa Indonesia and Acehnese. In other words, the speakers are from one ethnic group, Acehnese, from one Extending Circle country, Indonesia, who speak among themselves as part of an educational programme. This situation does not comply with the setting and characters specified by Jenkin; thus, the model may not apply to the speech community in the current study. The speech community in an international and academic conference, international schools and universities and business settings in Indonesia would suit this model better, which is not the focus of the current study.

Even though the setting and character of speakers in the current study may not be suitable for the EFL paradigm, some ideas carried out under this paradigm may help see the phenomenon in the current study more objectively. First, studies under this paradigm describe the findings in ELF as unique features instead of errors. For example, one study has documented salient features of ELF lexicogrammar, such as non-use of third person singular -s and pluralisation of uncountable nouns (Seidlhofer, 2004). Second, it encourages native speakers to adjust their perception and production when communicating with ELF speakers. It promotes that English textbooks should accommodate more ELF-oriented or NNE-oriented. Finally, it makes a very insightful critique of Second Language Acquisition (SLA) and English Language Teaching (ELT) scholars (Jenkins, 2006, 2012). ELF scholars reject interlanguage (a continuum between L1 and L2) and errors caused by L1 interference and fossilization (fixed error pattern in L2 learners' repertoire). These ideas all encourage scholars to look at the language development of English learners from a very different perspective away from what is

accepted as traditionally 'native speaker' and 'standard' oriented. Thus, ELF shares the same vision as the paradigm discussed in the previous section, the World Englishes paradigm, specifically the Three Circle Model.

2.3 Positioning English in Indonesia in relation to Neighbouring Countries.

This section first looks at the status of English and the features of English vowels in Indonesia's neighbouring countries, Malaysia, Singapore, and Brunei. Second, the history and development of English education in Indonesia are presented, followed by the vowel systems of Acehnese, Bahasa Indonesia, and English.

2.3.1 English in Indonesia's Neighbouring Countries

Instead of elaborating on English in the global context, it would be insightful to describe English in Indonesia in comparison to Englishes in neighbouring countries such as Singapore, Malaysia, and Brunei. These countries have immediate borders with Indonesia, and a possible encounter with the people from these countries is very likely. Malaysia and Singapore, in particular, are very close to Aceh and Acehnese often visit these countries for medical treatment and holidays. In addition, Singapore, Malaysia and Brunei share the Malay language as Indonesia (Kirkpatrick, 2010) but were colonized by different western countries. Brunei, Malaysia and Singapore were British colonies, and the status of English in these countries is very different from Indonesia. This historical background then shapes the status of English in these countries. Brunei and Malaysia, like Indonesia, opted to institutionalize Malay as their national language. However, like in Singapore, English remains important to Brunei and Malaysia but not to Indonesia. Singapore, on the other hand, in addition to Malay, also has three co-official languages, English, Mandarin and Tamil, to accommodate the multi-ethnic equity in the country (Ling, 2010). Schneider (2003) classified English in these countries as post-colonial English, while Kachru Kachru (1998) placed them in Outer Circle countries.

Malaysia

Malaysia maintained English along with Malay as its official language (Lowenberg, 1997) when Malaysia achieved its Independence in 1957 from the British. However, the government gradually favoured and promoted Malay over English during its development after independence. English medium schools were transformed into Malay medium schools by the Ministry of Education in 1969 (Platt & Weber, 1980) and by 1983, the process was completed and applied nationwide from primary to tertiary education. In addition to the Malay medium school, which is referred to as a national school, Malaysia also has a national type school in which the schools can use any of the three languages as the medium of instruction, Chinese, Tamil or English (Kirkpatrick, 2010). However, for a short period of time, between 2002 – 2012, Malaysia promoted English as a medium of instruction for science and maths subjects since Primary 1 but later reverted to Malay and vernacular languages (Kirkpatrick, 2010) because many students failed in science and math subjects.

Given the historical context of English in Malaysia and the current status of English in the country, it would be interesting to review the English variety of Malaysia, which is also known as Malaysian English (MalE). Malaysian English is characterized by the conflation of traditional English vowel pairs and the production of RP diphthongs as monophthongs (Pillai et al., 2010). It means that some Malaysian English vowels are similar, if not the same, in terms of quality and duration. Zuraidah (1997, pp. 38-40) found that Malaysians tend to produce the following vowel pairs as a single vowel: [iː] and [i] conflated to [i], [uː] and [ʊ] conflated to [u], [e] and [æ] conflated to [e], [ɒ] and [ɔ] conflated to [o], [ʌ] and [ɑː] conflated to [a], and [ə] and [ɜ] conflated to [ə]. However, the study was an impressionistic study that relied on hearing to judge the vowel production and are prone to subjectivity.

The later study by Subramaniam (2008) using acoustic analysis indicated that Malaysians indeed show some contrast in the traditional vowel pairs, but not as strong as those in RP. Despite the small number, the respondents of this study were fluent English speakers, which may contribute to the apparent vowel contrast produced. A more comprehensive study by Pillai et al. (2010) confirmed both studies to a certain extent. This recent study employing acoustic analysis similar to Subramaniam (2008) confirms the findings from the previous studies. It employed larger samples and was more inclusive and diverse because students from different backgrounds that is Malays, Chinese, Indians and Eurasian, participated in the study. In this study, Malaysian English was found to conflate the front vowel pairs [iː] - [i], [e] - [æ], [ʌ] - [ɑː] and exhibit lack of contrast for the back vowel pairs [uː] - [ʊ] and [ɒ] - [ɔː] (Pillai et al., 2010, p. 164). Similar to the findings in Singapore (Deterding, 2003) and Brunei (Sharbawi, 2006), Malaysian English also occupies a smaller vowel space compared to British English (Pillai et al., 2010).

Brunei

Brunei is an Islamic Sultanate country that immediately borders the Borneo Island of Indonesia. Together with Malaysian Muslims, the majority of its citizen share similar ethnicity, religion and culture with Indonesia. The Brunei population mostly comprises Brunei Malay (61%), Chinese (11%), and mixed Borneo indigenous groups and expatriates (20%) (McLellan, 2020, p. 399). Standard Malay and English is the official language of Brunei, while another local variety of Malay called Brunei Malay is used in an informal setting (Sharbawi, 2006, 2021). The status of English in Brunei is unique compared to Malaysia and Singapore. Compared to Malaysia, which abandoned the use of English as a medium of instruction in primary school, Kirkpatrick (2010) reported that Brunei has consistently used English as the medium of instruction since Primary 3. Recently, the new policy even introduced English starting in Primary 1 for science subjects while other general subjects such as Islam, Malay and Civics are taught in Malay

(Kirkpatrick, 2010). While Singapore was unsure which language to promote in its early development, Brunei is very confident in adopting English as the language of instruction and official settings. Even though English has gained popularity among its citizen, the country is unworried about the degradation of Brunei Malay as the people's "language of our soul" (Abdul Aziz, (1991) in McLellan (2020)) and the preferred code of communication in Brunei (Sharbawi, 2012).

With reference to Kachru's (Kachru, 1992a) Three Circle Model, Brunei is placed in the Outer Circle countries due to its protectorate history with the British Empire and its important role as a second language and language of education. Sharbawi (2012) placed Brunei in Schneider's (2003) English development model in Phase 3, *nativization*. This development seemed to be valid because the country is witnessing a shift of language from Standard Malay to English in the official setting. Ozog's (1996) study found that English was rated slightly higher by the Bruneians at 30.5 per cent compared to only 29.1 for Standard Malay. In a more recent study, Sharbawi (2012) also reported that English had gained important ground in education and media, formal and informal settings.

In terms of vowels, many features of Brunei English are similar to Singapore, the Philippines and Malaysia (McLellan, 2020, p. 403), but it also has features that are unique to Brunei. Previous studies have reported that Brunei English has compact vowel space (Sharbawi, 2006), exhibits rhoticity (Sharbawi & Deterding, 2010), and is inconsistence in the use of its vowels (Sharbawi, 2012). The most obvious features reported in Brunei English is rhoticity. Rhoticity is a pronunciation character in which the consonant /r/ is clearly articulated. In the rhotic accent, the orthographic *r* in words such as 'par' 'bird' and 'park' are produced. Considering Brunei was once a British colony, it is interesting to find its speakers to be rhotic, as rhoticity is often found in American English (Ladefoged & Johnson, 2014). Sharbawi and Deterding (2010) compared the

pronunciation of Singaporean and Bruneian reading the Wolf text and found that Bruneian tend to produce the text with r-coloring. The study found that around 50% of speakers from Brunei have a rhotic accent in NURSE and FORCE vowels. They proposed two reasons why Bruneians have a rhotic accent. First, Brunei Malay is a rhotic variety of Malay. So, there is a possibility that a transfer takes place from Brunei Malay to English among the Brunei speakers. Second is the possibility of the influence of American media. The sole provider of a satellite television network in the country is licensed by Kristal-Astro and air programs from the United States (McLellan, 2020).

Another notable feature of Bruneian English (BrunE) is its relatively compact vowel space. This was first reported by Sharbawi (2006). She found that even though the vowel of BrunE and Singapore English in the vowel plot seems similar, BurnE occupies a much smaller vowel space compared to British English. Brunei English speakers do not distinguish traditional vowel pairs to the same extent as Malaysian English and Singapore English. The vowel pairs [iː] - [ɪ], [e] - [æ], [ʌ] - [ɑː], [uː] - [ʊ] and [ɒ] - /ɔː/ are often conflated. However, [uː] - [ʊ] vowel pairs are much higher in Brunei English compared to Singapore English. For these pairs, Brunei English resembles more British English. The last feature in which recently reported by Sharbawi (2012) is that Brunei English seems to show inconsistencies in terms of pronunciation of English vowels. The last feature in which recently reported by Sharbawi (2012) is that Brunei English seems to show inconsistencies in terms of pronunciation of English vowels. The findings reported that there was inconsistent use of American English and British English. One of the examples she found was that some Bruneian speakers pronounce [ɑː] for either *bothered* or *flock* but never both (2012, p. 192). Another most salient feature is the influence of spelling on pronunciation (Sharbawi, 2012, p. 183). She found that some Bruneians pronounce the words, *company, concern* and *convinced* as LOT instead of STRUT, reading the vowel [o] as written.

Singapore

The majority of Singaporeans (70%) are of Chinese descent, (Kirkpatrick, 2010). On the other hand, Malay and Indians account for 14% and 8%, respectively, while the rest of the population are expatriates. In the past, The Singaporean Chinese were divided into Chinese-educated and English-educated. This would later result in the policy suppressing the spread of Mandarin while English domination gained momentum. As a result, in 2001, around 44% of primary school children reported that English was their home language (Jia & Fuse, 2007), and by 2009, the figure had reached 60% (Kirkpatrick, 2010). This indicates that English has become the first language for the majority of Singaporeans.

Singaporean success in adopting English as the main language would later result in the establishment of two varieties of English in Singapore, that is, the standard variety used in formal settings and the local variety used in informal settings and referred to by experts as Singapore Colloquial English (SCE) or Singapore English (Singlish) (Ling, 2010). She also reported that the local varieties of English in Singapore have some distinctive features in terms of vowel, lexical usage and syntax. While Singlish is referred to as Phase 3 by Schneider (2003), Deterding (2005, p. 194) suggests that Singlish appears to be in the fourth stage, which Schneider termed as "*endonormative stabilization*". Deterding (2005) also argued that Singaporean pronunciation is an emerging English style which is quite different from other known standard English and suggests that it should be described on its own instead of making a comparison to other known Englishes. Trudgill and Hannah (2013) classified Singaporean English as a second language variety of English and further explained how it is in the process of transforming from English as a Second Language variety to English as a Native Language.

It has been reported that Malaysian English and Singapore English are quite similar in that traditional vowel pairs such as [iː] - [ɪ], [e] - [æ], [ʌ] - [ɑː], [uː] - [ʊ], and [ɒ] – [ɔː] tend to merge into single phonemes. Ling (2010, p. 240) pointed out that Singlish does not maintain durational difference among the pairs, while Malaysian English sustains difference in duration except for [ɒ] - [ɔː]. Leimgruber (2011, p. 48) also argued that vowel length in Singlish is not contrastive enough to show a significant difference between the pairs consistently. Therefore, Singlish monophthongs are reduced to eight oral vowels consisting of three front vowels [i], [e], [ɛ], two central vowels [ə] and [ʌ], and two back vowels [u] and [ɔ]. This reduction in the vowels is similar to the Malaysian English reported by Zuraidah (1997), in which vowel pairs are often reduced into single vowels. In addition, Deterding (2003) also recorded that the [u] vowel in Singaporean English is a full-back vowel. Any words containing [uː] and [ʊ] vowels are more likely to conflate into [u] vowels and are located further back.

2.3.2 English and the Regional Language in Indonesia

Even though English is the only foreign language taught in Indonesia, the incorporation of English into the Indonesian education system experienced many alterations. The country has rectified numerous school education curricula. It battles between imposing Bahasa Indonesia as the national language, preserving its diverse local languages, and promoting English for a competitive global market.

Tracing back to when English was first pioneered in Indonesia, Horne (1961) claimed that it was introduced seven years after the Dutch colonies established an elementary school in 1907. During the Dutch colonial period, English was taught at schools but only limited to Dutch children. Very few Indigenous children got a chance to learn English, and those who did were mostly from privileged communities (Gregory, 1964). These

children eventually grew up learning Dutch and might have known some English, but it was not used as a means of communication.

Although the Netherlands occupied Indonesia for more than 350 years, the Dutch did not provide the Indonesian people access to education as much as did the United Kingdom to its colonies. Few Indonesians had access to education during this period, while only the Dutch children and a select few local people attended schools (Horne, 1961). Unlike the United Kingdom, which provided access to education for colonized communities, the Netherlands opted to let the colonized territories be uneducated. British colonies in adjacent areas such as Malaysia, Singapore and Hong Kong enjoyed the educational benefits brought upon by the British and eventually adopted English as their second language. Indonesia, occupied by the Dutch (Horne, 1961) and, despite the prolonged Dutch occupation, in fact, longer than any of those countries, refused to adopt the Dutch language.

In October 1928, during the Youth Pledge, the Indonesian youth from all over Indonesia came together to Jakarta and established a national identity by declaring Bahasa Indonesia as its national language. This identity declaration, coupled with the disgraceful attitudes of the Dutch toward their colonies, ignited identity awareness among the Indonesian throughout and kept the foreign languages away from disuse. As a result, neither Dutch nor English were widely spoken by the locals after the independence except for the few intellectuals who were closed to the Dutch. Horne (1961) pointed out that English instead of Dutch was adopted as a foreign language due to the negative image of the Dutch and the international status of English. Dutch policies were made such that even 15 years prior to its independence, around 94% of Indonesian was illiterate, and by 1940 only 30 senior high schools were reported throughout the colonized territories (Tilaar, 1995)

After independence, the Indonesian government was yet to decide which foreign language the country should use for international communication (Mistar, 2005) due to its historical western colonization and political turmoil in the early days independence. Only later did the Indonesian leaders choose English even though they had been educated in Dutch at schools (Huda, 1999) and Japanese for a short period. Compared to what happened in other British colonies in the regions, Indonesians were not prepared to adopt the language of the invaders (Thomas, 1968). However, English language teaching (ELT) did not take place until the Dutch acknowledged Indonesian sovereignty in 1949 (Mistar, 2005).

After the Dutch left the country, the Indonesian Ministry of Education and Culture (MEC) introduced an inspectorate of English Language Instruction and defined the objective of English in Indonesia, which is it "would never be either a social language or a second official language in Indonesia" (Sadtono, 1997). Later, as the number of Indonesians who participated in education increased, the MEC struggled to supply qualified English teachers. In 1950, to supply the high demand for English teachers, university students from any majors were recruited to teach in secondary schools (Sarumpaet, 1963) and existing teachers in the school were certified in a two-year evening course assisted by Ford Foundation (Mistar, 2005). The government also established programs to produce secondary and high school teachers in the same year, which were later integrated into the Teacher Training and Education Institute (*Institut Keguruan dan Ilmu Pendidikan-IKIP*) in 1954. This institute is the pioneer of teacher training and education faculties (*Fakultas Keguruan dan Ilmu Pendidikan*) which by 1961 were merged into universities across Indonesia (Mistar, 2005). Students who graduated from this faculty would later become teachers and teach school subjects at primary, secondary and high schools based on their subject qualification degrees.

The country has rectified numerous curricula and amended policies in school education as the government promotes national and local languages while embracing foreign languages. The Ministry of Education and Culture once introduced English in primary education before banning its teaching in 2013. Before introducing English to primary school, the government surveyed parents, teachers, and students on English teaching in Indonesia in 1985. It was found that even after six years of secondary and high school study, the English proficiency of high school graduates was far from satisfactory (Zein, 2017). The government believed that the absence of English in primary school might have inhibited their ability to successfully acquire English in secondary school (Sadtono, 1997). They were confident that the earlier English was introduced, the better the students would become. Then, in 1993 the MEC issued Decree No. 060 in 1993, which granted primary schools to teach English from Grade 4 if it was deemed necessary.

In early 2000, Indonesian parents became more aware of the need for English for their children, and more parents enrolled their children for English in elementary school (Zein, 2017). As a result, some schools even provide English as an extracurricular subject from Primary 1. To cope with the high demand for English at the primary level, the government gave the schools more freedom to accommodate the local need and demands by introducing the School-Based Curriculum, *Kurikulum Tingkat Satuan Terpadu (KTSP*). KTSP was stipulated in Decree No. 22/2006, allowing schools to teach English even from Primary 1 in the new curriculum. Learning English from primary school has made some parents and government stakeholders worry about losing local cultures and languages since school no longer gives room for this local subject. It has been reported that ten indigenous languages in Indonesia have vanished (Hadisantosa, 2010), and the number will continue to increase in the years to come. Thus, in 2013, MEC later introduced a new curriculum, the 2013 curriculum, also known as K-13. In this new curriculum, primary

schools can no longer teach English during school hours, and more time would be spent learning English at the secondary school (Zein, 2017).

Since English has gained popularity in these countries and is considered a must-have skill by many, some schools even promote the use of English as the language of instruction, as in the case of Bilingual Schools (Hadisantosa, 2010) and as the language of communication in Modern Islamic Boarding school in Indonesia. The function of English in Indonesian Islamic boarding schools seems to reflect the views of World Englishes views. The Islamic boarding school has its own unique English speech community to whom students interact with every day. Its usage is strictly limited to daily conversation within the boarding school with friends, teachers, and staff. Its establishment status is very much like the promotion of Bahasa Indonesia as the Indonesian national language in addition to the local language. Even though students can speak Bahasa Indonesia and local languages with their friends and teachers, they are prohibited from speaking those languages during school hours; they have to use English 24/7 during their three years in school.

The influence from both Bahasa Indonesia and the local languages in the students' English production is highly likely. One of such features may be apparent in their English vowel production. As for the Acehnese-Indonesian speakers in the present study, Acehnese and Bahasa Indonesia vowel systems may influence English vowels' production and contribute to their unique indigenized speech pattern. Thus, it would be interesting to see the features of English vowels produced by the Acehnese-Indonesian speakers of English at MIB schools. The following sections present the vowel inventories of Acehnese, Bahasa Indonesia, and English.

2.3.2.1 Bahasa Indonesia

Given the multilingual nature of Indonesia, that is, the fact that the majority of Indonesians may already speak more than one language before learning English cannot be ignored when one wishes to study how Indonesians acquire English. Bahasa Indonesia are taught English from the time they enter into secondary school as it is the only compulsory foreign language taught in the Indonesian education system (Kirkpatrick, 2012b; "Permendikbud No. 160," 2014). Thus, for most Indonesian, English is not a second language but a third language when they start learning it in secondary school.

The languages in which they are bilingual or multilingual may differ based on where they live and grow up in Indonesia. It is naïve to paint all Indonesian with the same colour as if Javanese/Indonesian and Acehnese/Indonesian bilinguals are the same and merely refer to them as Indonesian bilinguals. Even though most of the indigenous languages in Indonesia belong to the Austronesian family, their vowel system, for example, are likely to differ considerably. For example, Javanese, which belongs to the Malayo-Javanese language, only has six oral vowels while Acehnese being a family of the Malayo-Chamic language has ten vowels. Acehnese not only has a bigger vowel system compared to Javanese but also more vowels which are similar to English vowels.

It has been argued that when learning a second or third language, learners usually filter non-native vowels through their existing vowel system (Escudero, 2005). When one learns L2, L1 acts as the first initial state of their mental structure (Hummel, 2013) while when they acquire the third language, learners already have a combination of L1 and L2 phonemes in their mind (Amaro, 2012). Thus, with diverse local languages, Indonesian bilingual learners of English may acquire its vowel differently based on their indigenous language in addition to Bahasa Indonesia.

Bahasa Indonesia is the country's official and national language. Being a variety of the Malay language, Bahasa Indonesia had initially been used as the trading language in major urban areas in Indonesia (Moeliono, 1993), long before it gained independence. In 1918, the Dutch East Indies, which had colonized the Indonesian archipelagos for almost 330 years designated the language as its official language (Moeliono, 1993). Later in 1928, during the independence struggle against the Dutch, the youth from across the archipelago in Indonesia gathered and declared three commitments to unite all ethnic groups in Indonesia: to speak one language, accept one nation, and recognize one homeland called Indonesian (Ebing, 1997). When the Japanese annexed Indonesia from the Dutch, they promoted the use of Bahasa Indonesia in government settings and banned the use of any European languages to gain Indonesian sympathy in their pacific war. As a result, when proclaiming independence against the Japanese after their loss in the pacific war, the founding fathers of Indonesia spoke the language when broadcasting its independence to the world on 17 August 1945. The country then legitimised Bahasa Indonesia as its official language in the 1945 Constitution on 18 August of the same year.

The designation of Bahasa Indonesia as a national language in 1945 (Kirkpatrick, 2012b) has led to extensive adoption of the language by its citizens. Over decades, the number of Indonesians who claimed to speak BI as the first language had gradually increased from just 3% when it was first adopted in 1945 to 34% by 2000 (Montolalu & Suryadinata, 2007). It was also reported that the number of Indonesian speakers (both as a first and second language) was 60% in 1992 (Errington, 1992) and increased to 68% (140 million) by 2000 (Gordon). Quinn (2001), on the other hand, claimed that the number of Indonesians learning the language as their first language should be around 80% and rising by the 21st century. Thus, it was not a surprise when Hamers and Blanc (2008) mentioned that Indonesia together with India is among a few multilingual nations that were successfully able to impose a national language to its citizens where the national language

is used side by side with a local language. As a result, government success in diffusing BI as its official language has led to Indonesians becoming bilinguals in local languages and Bahasa Indonesia (Sneddon, 2003). Fishman (1978) even regarded Indonesia as a 'linguistic miracle' and 'the envy of the multilingual world'.

As the national language, Bahasa Indonesia is used in formal settings such as in schools, government agencies, businesses and media (Alwi & Sugono, 2000; Perwitasari, 2018), while local languages are used informally in the local market, traditional ceremonies and neighbourhood. Bahasa Indonesia is the official language of instruction in school and universities and all school subjects except English are written in Bahasa Indonesia. The current president of Indonesia, Jokowi, even encouraged Indonesian scholars to use Bahasa Indonesia as the official language at international conferences held in Indonesia. In government and business settings, Bahasa Indonesia is also used in formal letters, contracts, and government or business documents. Government officials use the language in public speech, bankers speak in Bahasa Indonesia to their customers, and both government and private employees use the language in their offices. Media such as newspapers, magazines, TV and radio also use Bahasa Indonesia. TV and radio broadcasters speak in Bahasa Indonesia, and both drama series and movies are also filmed in the language.

Acehnese/Indonesian bilinguals may acquire Bahasa Indonesia as their first or second language depending on where they live in Aceh, whether their parents speak Acehnese, and the attitude of the parents toward Acehnese. In the capital city, such as Banda Aceh, many people may learn Bahasa Indonesia as their mother tongue. They may only learn to speak or at least understand Acehnese only if their neighbourhood speaks Acehnese. Otherwise, they are monolingual in Bahasa Indonesia. Those who live in the outskirt of Banda Aceh or other districts in Aceh such as Nagan Raya, Aceh Barat and Aceh Jaya

may learn Acehnese as their first language at home from their parents and then start learning Bahasa Indonesia once they go to primary school. This is because people in rural areas use Acehnese more often in their daily conversation and have a more positive attitude toward the indigenous language.

At school, children may meet other kids who do not speak Acehnese, will learn Bahasa Indonesia as a school subject, and listen to teachers teaching school subjects in Bahasa Indonesia. During this primary school period, children who are monolingual in Acehnese will eventually become bilinguals in Acehnese and Bahasa Indonesia. On the other hand, children who are monolingual in Bahasa Indonesia may learn Acehnese if their friends speak to them in the language. In most cases, children's encounter with school education would create Acehnese/Indonesian bilinguals, Indonesian/Acehnese bilinguals, and Indonesian monolinguals who understand Acehnese. Indonesian who are bilinguals in other indigenous languages may have a different experience during their process to become bilinguals. Thus, when selecting Acehnese/Indonesian bilingual learners of English, it is important to consider this in mind.

Since multilingual speakers have more than one language repertoire, more factors are at play when learning new phonological system, size of current phonological system and closeness between the existing language and target language. On the one hand, Marx and Mehlhorn (2010) describe the mental state of bilinguals as having a bigger catalogue of phonetic-phonological parameters, increased language and metalinguistic awareness, and complex phonological knowledge. Due to their experiences, multilingual speakers benefit from their large langue repertoire, high awareness of metalinguistic features and help their acquisition of new language phonologies. All of these characteristics may facilitate or hinder learners' acquisition of L3 (Wrembel, 2015). On the other hand, Patihis, Oh, and Mogilner (2015) found that only learners whose L1 and L2 languages have similar

phonemic categories to that of L3 can facilitate the positive transfer. These findings confirmed an earlier study conducted by Beach, Burnham, and Kitamura (2001) who found that learners whose L1 and L2 are not similar to L3 could not get any positive benefit from their phonemic repertoire. Such facilitative and non-facilitative effect is often referred to cross linguistic influence (CLI) which may be influence by the closeness of existing and target language and proficiency of existing language (Westergaard et al., 2017). Based on this argument, Acehnese/Indonesian bilinguals may perceive and produce English vowels differently from the Javanese and Sundanese in (Perwitasari, 2019) study and might be able to benefit from their larger phonemic inventories with more vowel similar to English.

2.3.2.2 Acehnese Language

Aceh is one of the Indonesian provinces located on the tip of Sumatra Island. People from Aceh are often referred to in Acehnese as *Orang Aceh* 'Aceh People'. Acehnese is also the major ethnic group in the province, accounting for 90% of the total population in the Aceh province (McCulloch, 2005). In addition there are nine other local ethnic groups in this province: Alas, Gayo, Tamiang, Aneuk Jamèe, Kluet, Singkil, Simeulu and Haloban (Wildan, 2002). Gayonese, Alas, Tamiang, Aneuk Jame and Kluet are among the major ethnic groups after the Acehnese (Wu, 2012). Each ethnic group has their own language which they use among themselves. Despite Acehnese being the dominant ethnic group, the Acehnese language is not used as the lingua franca in the province. Instead, it is Bahasa Indonesia that is used for interethnic communication. Like the rest of Indonesia, people in Aceh learn Bahasa Indonesia in school as it is the national language.

The Acehnese language comprises various dialects. Asyik (1987) originally divided the Acehnese dialect into four major dialects, Banda Aceh, Pidie, Pase, and Meulaboh. The Banda Aceh dialect is the dialect spoken by the people living in Banda Aceh, the capital

city of Aceh and in Aceh Besar which is the district that encircles the capital city. Tanzir Masykar, Roni Agusmaniza, et al. (2021) argued that the Banda Aceh dialect should only be referred to the variety spoken in the capital city because it differs significantly from the Acehnese accent spoken in the Aceh Besar district. In fact, people living in Banda Aceh come from various regions in Aceh, and thus, speak Acehnese variably depending on where they come from. The Pidie dialect refers to the dialect spoken in what was previously the Pidie region situated in the north coast of Aceh. This region has now expanded into three regions: Sigli, Pidie Jaya and Meureudu. Another major dialect Asyik (1987) describes is the Pase dialect which is the dialect spoken in Pase in North Aceh. The Pase dialect has been previously claimed as "standard" Acehnese because there are not many variety within this dialect and a large number of speakers mentioned by Durie (1985). This claim needs to be treated with caution since there has been no official consensus made about which dialect should be considered as standard Acehnese. The Pase dialect is also the most heavily studied dialect of Acehnese (Asyik, 1987; Durie, 1985) which attracts the attention of linguistic scholars when talking about Acehnese.

The other major dialect described by Asyik (1987) was the Meulaboh dialect. Meulaboh is the capital city of the present-day West Aceh or Aceh Barat. Meulaboh used to be one region but is now divided into four regions: Aceh Jaya, Aceh Barat, Simeulu, and Nagan Raya. Simeulu is an island in the South of Aceh Barat. This island is mostly occupied by the Simeulu ethnic group who speak the Simeulu language. On the other hand, the population of Aceh Jaya, Aceh Barat and Nagan Raya are of Acehnese ethnicity and speak Acehnese. However, the people from these three districts speak different Acehnese dialects, with the Aceh Barat dialect being more dominant because it was the parent district prior to the division of the Aceh Barat districts into three more districts. However, people form Aceh Jaya, Aceh Barat and Nagan Raya share the same dialect feature unique to the Meulaboh dialect. For example, one distinguishing feature is the uvular trill in this

dialect while Acehnese's dialects in other districts use an alveolar trill. However, the former tends to be stigmatized as being 'vulgar' and 'impolite' in the way they use personal pronouns (Zulfadhli, 2015). This dialect is also an understudied dialect compared to other dialects such as Pase. By including this dialect instead of the other established dialects, it will enrich the literatures on diverse Acehnese dialects.

2.3.2.3 Acehnese Monophthongs

Acehnese has been reported to contain ten oral vowels and seven nasal vowels (Asyik, 1987; Yusuf, 2013). Both oral and nasal vowels have their monophthongs and diphthongs. The Acehnese oral vowels consist of three front vowels ([i], [e], and [ɛ]), four central vowels ([ɯ], [ə], [ʌ], and [a]) and three back vowels ([u], [o], and [ɔ]). The front vowel [i] is a close vowel while [e] is near-close vowel. The front vowel [ɛ] is a near-open vowel. The central vowels [ɯ] and [ə] are open and near-open, while the vowels [ʌ] and [a] are near-close and close, respectively. The back vowel [u] is an unrounded close vowel. The vowel back vowels [o] and [ɔ] are rounded near-close vowels and unrounded near-open vowels, respectively. The Acehnese oral monophthongs are presented in Table 2.1 taken from Asyik (1987).

Table 2.1: Acehnese Oral Monophthongs

	Front	Central	Back
High	[i]	[ɯ]	[u]
High-mid	[e]	[ə]	[o]
Low-mid	[ɛ]	[ʌ]	[ɔ]
Low		[a]	

In addition to oral vowels, Acehnese also has seven nasal vowels. All oral vowels except [e], [ə], and [o] have their nasal counterparts in Acehnese. They are the front nasal vowel

([ ] and [ɛ̃]), central nasal vowel ([ɯ̃], [ʌ̃], and [ ]) and back nasal vowel ([ ] and [ɔ̃]). The Acehnese nasal monophthongs can be observed in Table 2.2.

Table 2.2: Acehnese Nassal Monophthongs

	Front	Central	Back
High	[ ]	[ɯ̃]	[ ]
High-mid	[ɛ̃]	[ʌ̃]	[ɔ̃]
Low-mid		[ ]	
Low			

2.3.2.4 Bahasa Indonesia Monophthongs

There has been disagreement among scholars on the number of Bahasa Indonesia vowels. Previous works on Bahasa Indonesia vowels reported the number of vowels ranging from six to ten vowels (Dardjowidjojo, 2009; Lapoliwa, 1981; Marsono, 2008; Soderberg & Olson, 2008; Zanten, 1986; Zanten & Heuven, 1984). Different views on the allophonic variation of vowels primarily contributed to the disagreement. This difference in allophonic variation seems to depend on the background language of the speakers used in the study. Zanten (1986, p. 441) found that Bahasa Indonesia vowels were "influenced by the vowel system of their regional language when speaking Bahasa Indonesia and that allophonic variation depends largely on the regional background of the speaker."

Both Dardjowidjojo (2009) and Marsono (2008) argued that Javanese Indonesian has ten vowels, and this large number of vowels was contributed by Javanese vowels. Zanten and Heuven (1984) Lapoliwa (1981) Soderberg and Olson (2008) argued that Malay Indonesia (Zanten & Heuven, 1984) or sometimes referred to as Standard Indonesia (Soderberg & Olson, 2008) has six vowels ([i], [e], [ə], [a], [o], and [u]) and other variations are considered as an allophonic variation. The use of the phoneme [o] for the words cited in the previous study was also inaccurate, which seems to contribute to the confusion in the number of Bahasa Indonesia vowels. The vowel [o] was used for the

word '*bobot* [bobot] *weight*' when it should be represented with the vowel [ɔ], similar to the vowel [ɔː] in British English for the word port [pɔːt]. The vowel [o] can be found in words such as *bobok* [bɔbɔk] *sleep*, *bakso* [baksɔ] *meatball*, and *koko* [kɔkɔ] *male religious cloth*, while previous studies (Soderberg & Olson, 2008; Zanten & Heuven, 1984) used different words to capture the intended vowel. Lapoliwa (1981) argued that the vowel [ɔ] is the allophonic variation of /o/, and only appears in the final closed syllable. However, the [ɔ] vowel can be found in the final open syllable as in *toko* /tɔkɔ/ *shop*. Thus, the current study has used two different words to illicit both the vowels [ɔ] and [o] by the Acehnese speakers. The target Bahasa Indonesia monophthong vowels can be observed in Table 2.3.

Table 2.3: Bahasa Indonesia Monophthongs

	Front	Central	Back
High	[i]		[u]
High-mid	[e]	[ə]	[o]
Low-mid			[ɔ]
Low		[a]	

2.3.2.5 English Monophthongs

Choosing which English variety this study should refer to was not an easy task. As more people speak English, many varieties of English exist. According to Kachru (1992a), Indonesia is an expanding circle and a norm-dependent country which means that it has not yet developed its own variety. As a norm-dependent country (Kachru, 1992a), the English taught in Indonesian schools is subject to the norm provider countries such as the United Kingdom, the United States and Australia. It was reported that the first year high school English textbook used a mix of American and British English in its spelling Widiati et al. (2017). Lauder (2010) also previously stated that the Indonesian education

system has yet to decide which English varieties to use at school. However, since Hollywood movies are prominent on TV channels in Indonesia, American English was selected as the basis of English vowels for the current study. The English monophthongs can be observed in Table 2.4.

Table 2.4: English Monophthongs (Ladefoged & Johnson, 2014, p.96)

	Front	Central	Back
High	[iː] [ɪ]		[uː] [ʊ]
High-mid	[e]	[ʌ]	[ɔː]
Low-mid	[æ]	[ɝ]	
Low			[ɑː]

According to Ladefoged and Johnson (2014), Standard American Newscaster English has a slightly different vowel compared to British Broadcasting Corporation (BBC) English. Standard American English is the English variety spoken in the United States of America and Canada. American English has 16 vowels consisting of nine monophthongs, six diphthongs, and one rhotic vowel. The nine monophthongs and one rhotic vowel can be observed in Table 2.4

2.3.2.6 English vowel contrast

Different studies have used different English vowel pairs to highlight the difference in production and perception of English vowel contrast. The various pairs used were mainly determined by the English variety used as the reference. Studies with reference to British English variety under the World Englishes paradigm employed five English vowel

contrasts, [ɪ] - [iː], [e] - [æ], [ʌ] - [ɑː], [ɒ] - [ɔː], [ʊ]- [uː] (Pillai, 2014; Pillai et al., 2010). On the other hand, studies conducted by using the same British English varieties as reference under the same World Englishes paradigm, only compared four English vowel contrasts (Deterding, 2005; Sharbawi, 2006). Sharbawi (2006) conducted the study on the English vowel production by Bruneian speakers in which British English was used as the basis for the study. This study looked at four English vowel contrasts [ɪ] - [iː], [e] - [æ], [ɒ] - [ɔː], [ʊ]- [uː] and did not include the [ʌ] - [ɑː] in her study. Since the reference for the study was British English, citing Deterding (1997) study, she argued that British English speakers commonly produced [ʌ] - [ɑː] with almost the same quality and only differed in terms of duration. It must be noted that Sharbawi (2006) was a PhD student under the supervision of Deterding when she conducted the study which explains Deterding's influence on her selection of English vowel contrast.

The study by Deterding (2005) on Singaporean English was among the first studies describing Singaporean English as an emergent English variety. Similar to Malaysian English and Bruneian English, British English was also the basis for Singaporean English. However, the study only plotted three English vowel contrasts by the Singaporean speakers [ɪ] - [iː], [e] - [æ], [ɒ] - [ɔː]. In addition, excluding the [ʌ] - [ɑː] pair in the plot of English vowel contrast, the study also did not plot the [ʊ]- [uː] pair. Deterding (2005) found that the quality difference for the [ʊ]- [uː] pair among the Singaporean speakers were bigger than the British English speakers. However, since the number of the speakers was small, he argued that the finding may not be representative of the real production of the [ʊ]- [uː] pair, suggesting for further study with a larger number of speakers. The study by Low (2016) which also looked at Singaporean English also employed five English vowel contrasts similar to Pillai (2014). She opted for five English vowel pairs because she included speakers from other countries in her study, that is India, Chinese, Philippines, and Indonesia.

In terms of studies on Acehnese, Fata et al. (2017) found that Acehnese Indonesian high school students did not separate four English vowel contrasts in their production of [ɪ] - [iː], [e] - [æ], [ɒ] - [ɔː], [ʊ]- [uː]. This study did not mention whether British English or American English was used as the reference. The study also found that the [ʌ] - [ɑː] pair was produced with a distinctive quality by male speakers. Female speakers on the other hand produced [ʌ] - [ɑː] at a rather close distance. Since no statistical analysis was conducted, it was unclear if the male and female speakers produced this pair significantly distinctively. Thus, the current study included this pair in the English vowel contrast analysis to confirm this assumption.

As for American English, the five English pairs may use different phonetic symbols especially for the British English vowel [ɒ]. American English does not have the vowel [ɒ] because words containing the British English [ɒ] actually are produced as [ɑː]. Thus, the vowels in the words such as *pot* and *bard* are both produced as [ɑː] in British English. Since the current study used American English as the reference, the American English [ɑː] is used for [ɑː] - [ɔː]. Another difference worth pointing out is the symbol used for the [e] vowel. Roach (2009) who described American English used the [e] symbol for the vowel in *beg* and *bet* while Ladefoged and Johnson (2014) opted for the [ɛ] symbol for describing the vowel in both American and British English. Since Ladefoged and Johnson (2014) included both American and British English for the [ɛ] symbol, the [ɛ] is used for the [ɛ] - [æ] pair. Thus, the five English vowel pairs used in the current study are: [ɪ] - [iː], [ɛ] - [æ], [ɑː] - [ɔː], [ʊ]- [uː]. For further presentation of words used to collect the perception and production data, see 3.4 on data collection.

2.4 Studies on English Vowel Production by Indonesian Speakers

There have been numerous studies on English vowels produced by Indonesian speakers. Some studies focused on vowel qualities (Fata et al., 2017; Low, 2016; Perwitasari, 2019;

Perwitasari et al., 2016; Widagsa, 2015; Widagsa & Putro, 2017); others looked at the durational length (Perwitasari, 2019; Perwitasari et al., 2015). Most studies on English vowel qualities by Indonesian speakers compared them to more established varieties such as British English (Widagsa, 2015; Widagsa & Putro, 2017) or American English (Perwitasari, 2018, 2019; Perwitasari et al., 2016). One study (Low, 2016) compared the quality of English production by Indonesian speakers to other emergent varieties such as Singapore and Malaysia, while another study by Fata et al. (2017) described English vowel quality without any comparison to different English varieties.

A group of previous studies has attempted to focus on vowel qualities by non-Acehnese speakers (Perwitasari et al., 2016; Perwitasari et al., 2015; Widagsa, 2015; Widagsa & Putro, 2017). To begin with, both Widagsa (2015) and Widagsa and Putro (2017) looked at the production of English vowels by undergraduate university students from various parts of Indonesia. The students were English Department students at the PGRI University in Yogyakarta who have studied English for three years. In both studies, the authors found that the speakers tended to produce some English vowels similarly, [ɪ] - [iː], [ʊ]- [uː], and [e] - [æ]. In addition, Widagsa and Putro (2017) also compared English vowels produced by Indonesian speakers to British English vowel quality. The quality of English vowels used as the basis of comparison was taken from another study by Hawkins and Midgley (2005) on British English Received Pronunciation. They argued that Indonesian speakers tended to have a smaller vowel space compared to British English speakers.

The study by Widagsa (2015) further attempted to compare the quality of English vowel production to the quality of Bahasa Indonesia vowels. Based on their comparison of English vowel production to Bahasa Indonesia vowels, Widagsa (2015) concluded that Indonesian speakers produced English vowels with approximation to existing Bahasa Indonesia vowels, which lends its support to the Perceptual Magnet Hypothesis by

Iverson and Kuhl (1995). The [ɪ] – [iː], [ʊ] – [uː], and [e] – [æ] pairs were produced in close proximity to Indonesian [i], [u] and [e]. However, the quality of Indonesian vowels used as the basis for comparison was taken from another study by Dardjowidjojo (2009) instead of recording the actual production of Indonesian vowels produced by the same speakers in the studies.

Subsequently, Perwitasari et al. (2016) compared the quality of English vowels produced by Javanese and Sundanese speakers to American English vowels produced by American speakers living in Indonesia. The Javanese and Sundanese speakers were university students from Central and West Java majoring in various fileds from several universities in Yogyakarta. Perwitasari et al. (2016) found that Javanese and Sundanese speakers tend to have a smaller vowel space compared to American speakers. The study argued that having a smaller vowel space, Javanese and Sundanese had difficulties producing English vowels which have more crowded vowels. The Javanese and Sundanese speakers also had difficulties producing English vowels absent from the Indonesian vowel system ([iː], [æ], [ɑː], [ɔː], [uː], [ʌ], [ɜː]). Earlier, Perwitasari et al. (2015) analysed the durational length of the English vowels produced by Javanese speakers and compared them to American speakers. The study found that Javanese speakers failed to realize both English short and long vowels and generally produced every vowel shorter than the native American speakers.

In another study, Low (2016) recorded English vowels produced by one Indonesian speaker who was a PhD student in Singapore. Instead of comparing it to British or American English, she compared the production by Indonesian speakers to Chinese, Filipino, Indian and Singaporean speakers. She found that the Indonesian speaker was able to distinguish typical vowel pairs. The Indonesian speaker's production of the [ɪ] – [iː] vowels were rather close, similar to Filipino and Chinese speakers, but not conflated

as in the Singaporean speaker. The speaker was also able to distinguish the [ʊ] – [uː] vowels but produced the [ʌ] – [ɑː] in proximity. It seems that the English developmental stage of the speaker may contribute to the speaker's ability to distinguish most typical vowels compared to the results found in other studies (Perwitasari et al., 2016; Widagsa & Putro, 2017). Despite generating exciting results, this study only recorded one Indonesian speaker who may not represent the diverse ethnicities existing in Indonesia. No information was given on whether the speaker speaks a vernacular language in addition to Indonesian. Therefore, it is unclear which part of Indonesia the speaker represents.

In terms of theoretical backgrounds, the studies by Perwitasari et al. (2016), Widagsa (2015), and Widagsa and Putro (2017) all based their assumption on theories of second language acquisition: Speech Learning Model (SLM) by Flege (1995) in Perwitasari et al. (2016) study, Feature Hypothesis McAllister et al. (2002) in Perwitasari et al. (2015)and Perceptual Magnet Hypothesis by Iverson and Kuhl (1995) in Widagsa (2015), and Widagsa and Putro (2017) studies. Only Low (2016) used the World Englishes paradigm. By referring to SLA theories, these studies treat English vowels produced by Indonesian speakers as deviant from established English varieties such as British and American English. Such comparison diminishes the possibility of observing English vowels produced by Indonesian speakers as emergent English varieties similar to other emergent varieties in adjacent countries such as Singapore, Malaysia and Brunei. Given Indonesia's immediate border with these countries, it would be more reasonable to compare English vowels produced by Indonesian speakers to English varieties in adjacent countries.

Moreover, to date, only one publication focused on vowel production by Acehnese speakers. Fata et al. (2017) recorded the production of English vowels by 10 Acehnese-

Indonesian high school students. They found that the Acehnese students produced the [ɪ] – [iː], [e] – [æ], [ʊ] – [uː], [ɑː] – [ɔː] closely together. The results are similar to previous studies on Javanese and Sundanese, except that the Acehnese speakers produced the vowel [ɑː] closer to [ɔː] while the Javanese and Sundanese produced [ɑː] closer to [ɔː]. The study by Fata et al. (2017) is the first study that attempted to describe English vowels produced by Acehnese speakers and the first study on Indonesian speakers recording many participants that did not make the comparison to the established English varieties. However, there are a few shortcomings found in this study. First, the speakers who participated in the study were from conventional high schools who only learned English during the English subject in school. Thus, their use of English may be limited to the classroom context. Second, no information on student English proficiency was provided, so it was unclear what their English proficiency level was. Third, it was also unclear on which paradigm it refers to when describing the English vowels produced by Acehnese-Indonesian speakers. Even though the study does not make any comparison to any existing English varieties, the study by Fata et al. (2017) mostly cited studies that use SLA theories such as SLM by Fox et al. (1995) and Second Language Linguistic Perception (L2LP) by Escudero and Williams (2011). On the other hand, most studies on production cited were studies situated under the World Englishes paradigm, such as Malaysian English (Pillai, 2014), Singaporean English (Deterding, 2005) and Brunei English (Sharbawi, 2006).

Thus, the present study examines English monophthongs produced by Acehnese speakers in terms of vowel quality and length contrast. Using the World Englishes paradigm, this study does not intend to make any comparison to the "standard" English varieties such as American and British English. Instead, this study describes Acehnese Indonesian speakers' production of English monophthongs as a possible emergent English variety.

In addition, these previous studies also viewed the data under the SLA paradigm instead of Third Language Acquisition (TLA) paradigm. In SLA, Cross Linguistic Influence (CLI) is seen as one to one influence from L1 to L2, but in TLA, CLI is seen as a combined influence of previously acquired languages on the target language (De Angelis, 2007). In other words, when learning a third language, knowledge obtained when acquiring a first and second language influences the third language acquisition. Considering the multilingualism context of Indonesia, the influence of the first and second languages on the production of English vowels is inevitable. However, previous studies have produced mixed results on which languages are more influential when learning L3. Some studies suggested that L1 had a greater influence on L2 (Wrembel, 2012) while others suggested that L2 has a strong influence on L3 at the initial stage of acquisition (Wrembel, 2013). A previous study also reported the combined influence of L1 and L2 on L3 (Lipińska, 2015).

Ringbom (1987) claimed that foreign language accents are strongly influenced by L1 regardless of L3 proficiency. He also argued that the influence of L2 on L3 phonology is rare and depends mostly on the intensity of L2 usage. Supporting this claim, Llisteri and Poch (1987) found that Catalan/Castilian bilinguals learning French exhibited no interference from L2 in the production of French vowels and fricative consonants. However, a more recent study by Wrembel (2012) showed some support for L1 dominant influence on L3 phonology. The study found that irrespective of the level of L3 language proficiency, the influence of the LI Polish instead of L2 French on the L3 English phonology prevailed in their accented speech. In addition, influence from L2 French to L3 English was also recorded but to a lesser extent. In a later study, Wrembel (2013) investigated the production of L3 French by L1 Polish and L2 English speakers. Despite yielding the strong influence of L1 Polish on the production of L3 English, L2 influence

was found to be considerable. The study later suggested that combined cross-linguistic influence may take place in bilinguals learning L3.

However, it is important to point out that these studies (Wrembel, 2012, 2013) tend to employ foreign accent ratters in determining the influence of L1 and L2 on L3 phonology. The studies also did not specifically account for cross-linguistic influence at phonemic levels. One of the recent studies that looked at the influence of L1 and L2 on L3 phonology at the phonemic level was conducted by Lipińska (2015). Instead of employing foreign accent rating, the study compared the formant values of L3 German [œ] to L1 Polish [u], [ɔ], and [ɛ], and the English [uː] and [ɔː]. The German [œ] was selected because Polish learners of German found it difficult to produce this vowel. Interestingly, even though the German [œ] was assumed to be closer to the Polish [ɛ], its F1 values were too low pushing it much closer to Polish [u] and English [uː]. Some speakers also produced the German [œ] closer to the Polish [ɔ] and the English [ɔː]. The results suggest that there is a combined influence of L1 Polish and L2 English on the production of the German [œ].

All of the mentioned studies compared the production of L3 by the speakers in the European context in which most languages may share the same family root with the Indo-European family. In the South East Asian context, Yunus and Pillai (2020) recorded the production of L3 German vowels by 10 female speakers of L1 Malay and L2 English. The study specifically compared the quality of German vowels to Malay and English vowels in order to observe CLI from Malay and English on German vowels. The study recorded a CLI effect on the production of German vowels. The influence mostly came from L1 Malay instead of L2 English. Six German vowels ([ɪ], [eː], [ʊ], [uː], [ɔ] dan [a]) were produced with the same quality to five Malay vowels ([i], [e], [u], [o] dan [a]). Furthermore, L3 German [ɪ] was produced with the same quality and duration as L1

Malay instead of L2 English. The speakers' L1 is Malay and shares the same root as Indonesian, a Malayo-Polynesian language family (Britannica, 2013). Thus, similar findings may be expected from English vowels produced by Indonesian speakers. However, L3 English and L2 German investigated in this study belong to the Germanic group of Indo-European languages (Britannica, 2020). In the case of Indonesia, Indonesian and Acehnese belong to the same Malayo-Polynesian language family. Therefore, both Bahasa Indonesia and Acehnese may simultaneously influence the production of English vowels by Acehnese-Indonesia speakers.

However, in the Indonesian context, no studies have compared English vowels with both Indonesian and the local language vowels produced by the same speakers. One study by Widagsa (2015) compared the production of English vowels by Indonesian speakers to Indonesian vowels, but the Indonesian vowels were taken from another study. This comparison does not reflect the actual influence of Indonesian vowels on English vowel production. Furthermore, the speakers who participated in the study were from various parts of Indonesia, even though the authors admitted that Indonesians might produce Indonesian vowels differently across regions (Widagsa, 2015; Widagsa & Putro, 2017). Even though Perwitasari (2019) compared English vowels to Sundanese and Javanese vowels, the study did not account for Bahas Indonesia vowels and dismissed it as the subset of Javanese and Sundanese vowels. She argued that Sundanese and Javanese have more vowels than Indonesian, and all vowels in Indonesian exist in Javanese and Sundanese. Thus, in both studies, the multilingual nature of Indonesian speakers was not considered.

2.5 Models of Speech Perception

Before embarking on the studies of vowel perception in Indonesia, this section first review existing models in speech perception of non-native sounds. Studies on the

perception of non-native sounds (L2, L3 or additional language) are trying to explain how language learners perceive non-native sounds in relation to the sounds in their existing language (Escudero, 2005). Many models have been proposed to explain how these existing sounds affect the acquisition of non-native sounds, like SLM (Speech Learning Model) (Flege, 1995), FBM (Feature-based Model) (Brown, 1998), L2LP (Second Language Linguistic Perception) (Escudero, 2005), and PAM (Perceptual Assimilation Model) (Best, 1994; Best & Tyler, 2007).

Barrios et al. (2016) argued that these models could be divided into two different approaches of similarity in cognitive science, spatial approach and featural approach. In the featural approach, sounds are viewed as having a set of discrete features and their similarity and differences are determined by these features. FMB (Brown, 2000; Brown, 1998) falls under this category. In the spatial approach, sounds are viewed as "objects represented as points in a continuous, multidimensional space" (Barrios et al., 2016, p. 368). In this approach, the similarity between sounds is determined by how close one sound is to another in the multidimensional space. The three most influential models in language perception, that is SLM (Flege, 1987, 1995), PAM (Best, 1994; Best & Tyler, 2007) and L2LP (Escudero, 2005), fall under this category.

According to FBM (Brown, 2000; Brown, 1998), 'the learner's native grammar constraints which non-native contrasts he or she will be able to accurately perceive and, therefore, limits which non-native contrasts the learner will successfully acquire' Brown (2000, p. 19). In other words, learners can only perceive and produce non-native contrasts if their existing language has the features required to perceive and produce the non-native contrasts. This process is called redeployment a term coined by Archibald (2005) in which learners redeploy their L1 features when acquiring L2 sounds. In support of her claim, Brown (1998) compared the ability of Japanese and Chinese learners of English to

discriminate the English [ɹ] – [l] contrast presented in AX test. The study found that Japanese learners of English failed to distinguish minimal pairs involving English [ɹ] – [l] contrast while Chinese learners did not have any difficulties distinguishing these contrasts. Based on this finding, she concluded that the Japanese learners failed to realize the English [ɹ] – [l] contrast because [l] is absent from Japanese, while Chinese were able to redeploy their L1 [ɹ] – [l] to distinguish the English [ɹ] – [l] contrasts.

However, a growing body of literature has shown that L2 learners were able to produce sounds whose features were absent from their L1. One such recent study in vowel contrast was conducted by Barrios et al. (2016), comparing the perception of the English [ɑ] – [æ] and [i] – [ɪ] by advanced Spanish learners of English. Both [ɑ] – [æ] and [i] – [ɪ] were reported to be problematic for Spanish learners of English. In both pairs, the vowel [æ] and [ɪ] does do not exist in the Spanish vowel system. Based on feature analysis, in the [ɑ] – [æ] pair, the vowel [ɑ] is a low back vowel while [æ] is a low front vowel, while in the [i] – [ɪ] pairs, both vowels are high front vowels. Brown's featural model predicted that Spanish learners would be able to redeploy the [ɑ] feature from Spanish to distinguish it from the non-native English [æ]. As for the [i] – [ɪ], the Spanish learners of English would fail to perceive the difference in the two sounds because of features required to distinguish the low front vowel pairs [i] – [ɪ] (both quality and length) which were absent in Spanish. Contrary to the prediction, the advanced Spanish learners of English were equally successful in discriminating both [ɑ] – [æ] and [i] – [ɪ] presented in both AX and ABX perception tests.

Moving to spatial approach, there are three most influential models in language perception, that is, SLM, PAM and L2LP. The three models differ in their aim, proposition and prediction regarding the perception of non-native sounds in relation to native sounds. First, SLM presented by Flege (1987, 1995) aimed to explain the ultimate

attainment of language learners in achieving native-like pronunciation. Studies conducted by Flege and colleagues under this model observed migrants or language learners acquiring a second language who live in the target language environment (Flege et al., 1999; Flege & MacKay, 2004; Flege et al., 2003; Flege et al., 1995). The core idea of SLM is that the authentic production of L2 sounds often has a perceptual basis (Flege et al., 1995). When learning L2 sounds, learners identify the L2 sounds through their existing L1 sounds which they have established during L1 acquisition (Flege, 1987). Flege (1991) called this process of language processing as "interlanguage identification" and compared the process as "pegging square pegs into round hole" or "pouring new wine into old bottles." In relation to this, Flege (1995) made seven hypotheses regarding the perception and production of second language sounds. Only three hypotheses, H2, H3 and H7 (Flege, 1995, p. 243) are discussed in relation to the current study. According to H1, if L1 has a smaller vowel system than L2, "L2 learners are likely to discern the phonetic differences between certain L1 and L2 vowels." In relation to this, Flege argued that Spanish speakers should be able to perceive English vowels more easily because Spanish has fewer vowels than English. In H2, SLM stated the greater the difference between the L2 sound and the closest L1 sound is perceived, the more likely phonetic differences are discerned. An example of this is the Spanish speaker's perception of the English [æ] and [iː]. The model predicts that Spanish speakers may likely be able to establish the phonetic category for the English [æ] compared to the English [iː] because [iː] differs only slightly from the Spanish [i]. In H7, Flege (1995) argued that the production of the L2 sound eventually resembles the perception of the said sound. This last hypothesis suggests that L2 learners may fail to distinctively produce L2 vowel contrasts if they could not discriminate the contrasts in the perception test.

Based on phonemic similarity between L1 and L2, Flege (1987) taxonomically divided L2 sounds into three distinctive categories, identical, similar and new. These categories

are believed to represent how L2 learners perceived non-native sounds in relation to their L1 sounds. It has been found that L2 learners' representation of the non-native sounds differs from that of monolingual speakers (Flege, 1991). Matching L2 sounds with L1 sounds to determine their categories is not always straightforward because the exact quality of sounds in each language will always be different despite being produced similarly (Flege, 1987). This comparison can be made through phonemic and phonetic comparisons. The phonemic comparison involves only the universal representation of certain sounds by referring to the IPA symbol. English vowels in the current study would be divided based on the classification model proposed by Flege (1987) and Flege (1991). If two sounds in two different languages being compared share the same IPA symbol, the two sounds are considered identical. If the two sounds in the two languages being compared share the same symbol but one of the sounds has an additional feature such as tense with the [ː] symbol while the other language does not, the two sounds are classified as similar. If the two sounds in the two different languages do not share the same IPA symbol and additional features, the two sounds are classified as new. However, the term identical needs to be treated with caution.

Second, the Perceptual Assimilation Model (PAM) explains the acquisition of non-native sounds at the initial stage of language acquisition as seen in Best (1994); Best and Tyler (2007). Initially, the model (Best, 1994) was developed to describe the way infants perceive new sounds in their L1 environment. Later, the model was extended to account for the acquisition of a second language (Best & Tyler, 2007). PAM posits that when non-native listeners listen to non-native sounds, they will assimilate the sound to any existing similar L1 sounds due to their experience of their native sounds. The listeners will assimilate the non-native sounds into three possible categories based on their goodness fit to native sounds. They will hear the sounds as a good or poor example of their L1 sounds (categorized), distinct from any of their L1 sounds (uncategorized) and as ambient

non-speech sounds (non-assimilated). Studies on PAM employed non-native contrast to determine the level of difficulty of perception. Based on the non-native contrasts, the categorized assimilation model could be divided into three possible scenarios, and their ease of acquisition is based on these scenarios. If two non-native contrasts are assimilated into two different native sounds, the assimilation is called a *two-category* scenario. The *two-category* scenario is deemed to be easy to discriminate with because the non-native contrasts are assimilated into two comparable native contrasts. However, the discrimination would be less accurate when the two non-native contrasts are assimilated into a single native phoneme, either as an equally good or poor example of the native phoneme. This type of assimilation is called a *single category*. Discrimination varies from good to very good if one of two non-native contrasts is perceived as a member of one L1 while the other is not. This type of assimilation is named *category-goodness*. Category goodness has two additional possible assimilations. Discrimination could be good if one of the contrasts is assimilated into the native sound (*categorized*) while the other is heard as a non-native sound (*uncategorized*). This combination is called (*categorized-uncategorized*). Discrimination would be moderate if both sounds are heard as non-native sounds and are not assimilated to any existing L2 sounds (*uncategorized-uncategorized*). The last hypothesis, which rarely happens, is when the contrasts are non-assimilable because the sounds are heard as non-speech sounds. Discrimination can be poor or excellent depending on whether the participants discern the difference between the non-speech sounds.

Lastly, the L2LP model aimed to accommodate what was left by the previous two models (SLM and PAM). The model was proposed by Escudero (2005) in her PhD thesis. According to her, PAM only accounted for the perception and production of non-native sounds at the initial stage of language development. At the same time, SLM accommodated the perception and production of non-native sounds at the end stage of

language development. Escudero (2005) claimed that none of the two models accounted for the developmental stage of second language acquisition. Thus, L2LP was proposed to explain the perception and production of non-native sounds at the initial, developmental and end stages of second language acquisition. To explain the cross-linguistic influence of L1 and L2, the model expects that optimal perception of L1 and L2 sounds is first established. Optimal perception is defined as the maximum-likelihood condition a person perceives of a certain sound as the intended sound in the listener's production environment. For example, if the vowel [i] is produced at the F1 value of 280Hz and no other sound in this language is produced at this F1 value, it is 100% likely that this F1 value will be perceived as the front vowel [i] in the said language. Thus, when determining the ease of acquisition of a target sound, the F1 and F2 values of the target sound are first determined as the optimal condition of the sound. When these values are compared with the optimal condition of the closest sound in speakers' L1, the optimal condition L2 acquisition can be established. The optimal condition of this L2 sound depends on the vowel system of speakers' L1 and target language and the ability of the learners to perceive the distance between the optimal L1 condition and optimal L2 condition.

Regarding this cross-linguistic comparison, the model made three possible scenarios and predictions of non-native vowel perception: *new*, *subset* and *similar*. The *new* scenario occurs when two non-native sounds are perceived as a single sound. This scenario is considered very challenging for L2 learners because they are copying their L1 category to perceive the non-native sounds. Their perception of the L2 sound is equal to their perception of L1 sound. This full copying of L1 is expected to occur among beginner L2 learners. For example, the early Spanish learners of English initially perceive the English vowel contrast [iː] – [ɪ] in the word sheep – ship as the Spanish [i]. The *subset* scenario occurs when two non-native sounds are perceived as more than two existing L1 sounds.

This scenario mostly happens if the target L2 sounds constitute the subset of L1 sounds. An example of this scenario is that Dutch learners of Spanish perceiving the Spanish [i] – [e] as the Dutch [iː], [ɪ] and [ɛ]. This scenario also poses a challenging problem because the learners may perceive that more vowel contrast exists in Spanish. The last scenario is a *similar* scenario in which L2 learners equate two non-native sounds to two existing native sounds. This scenario is predicted to be easier compared to the other two scenarios (new and subset). An example of this scenario is that Spanish learners of English equate English [iː] – [ɪ] contrast into Spanish [i] – [e].

2.6 Studies on the English Vowel Perception by Indonesian Speakers

In regards to the Indonesian context, there are two recent studies on the perception of English vowels by Indonesian speakers of English (Perwitasari, 2013, 2018). Perwitasari (2013) compared the perception of English vowel contrasts using context and non-context listening tests. The study found that the speakers produced significantly fewer errors when discriminating vowel contrast in context than without context. The error rates of some vowel contrast for the non-context-based test reached 50% for [iː] – [ɪ]. The figures were mostly below 30% in the context-based test except for [iː] – [ɪ] at 35% for one of the high front vowel pairs. However, it was unclear in the study what constituted context and non-context-based tests because the study did not clearly indicate how the stimuli provided in the context test were different from the non-context test. The study also did not mention the number of Indonesian speakers participating in the test. Instead of viewing the data as learners' ability to distinguish English vowel pairs, the study interpreted the data as errors. Thus, the misperception score of 25% and below on certain vowel contrasts were viewed as errors in the study instead of 75% accuracy. However, according to Jia et al. (2006) a score of 70% accuracy in the perception test is considered above chance level and should be categorized as good perception ability. Even the English monolinguals in

the study by Jia et al. (2006) performed at 70 – 80 percent of accuracy for some vowel contrasts.

In the later study, Perwitasari (2018) examined the perception of English vowel contrasts by Javanese and Sundanese speakers of English. Specifically, the study determined the ease of perception of English vowel contrast based on the acoustic similarity and differences between the Javanese and Sundanese vowels and English vowels. The study divided English vowels into new and similar categories following Flege (1987) classification of L2 sounds against L1. New vowels were English vowels absent from Javanese and Sundanese vowel systems. The new vowels included [ɑː], [ɜː], [ɔː], [ʌ], [æ], [ɛ], [ɪ], and [ʊ]. Similar vowels were English vowels that use the same IPA symbol as Javanese and Sundanese vowels but have a longer duration. Two vowels were categorized as similar vowels, [iː] and [uː]. In this study, the Javanese and Sundanese speakers (30 each, 15 males and 15 females) and American speakers (20, 10 males and 10 females) participated in the perception of English vowel contrasts using the mouse tracking study. The study found that both Javanese and Sundanese speakers had lower perception accuracy for new vowels supporting the L2LP model (Escudero, 2005) and rejecting the SLM model (Flege, 1995). As for similar vowels, the Javanese speakers performed at a lower accuracy for the [uː] vowel, while Sundanese had lower accuracy for [iː] and [uː]. The perception accuracy scores were higher in American speakers than Javanese and Sundanese speakers in all vowels tested.

In the context of Acehnese speakers, to the best knowledge of the researcher, there has been no study on the perception of English vowel contrast. Compared to Javanese and Sundanese vowels, Acehnese has more vowels and may redeploy their existing vowels differently when perceiving English vowel contrasts. For example, the English [ɔː] and [ʌ] were considered new vowels in the Javanese and Sundanese vowel systems, according

to Perwitasari (2018). For Acehnese speakers, the English [ɔː] can be classified as a similar vowel, while the English [ʌ] is identical. Acehnese has the vowel [ɔː] but without durational length, while the English [ʌ] is identical to the Acehnese [ʌ]. There was no English vowel in Perwitasari (2018) identified as identical to Javanese and Sundanese. This additional classification might provide different results regarding perception accuracy of English vowel contrasts by Acehnese speakers. Thus, for Acehnese speakers, the new vowels are [ɑː], [ɜː], [æ], [ɛ], [ɪ], and [ʊ], similar vowels are [ɔː], [iː] and [uː] and identical vowel is [ʌ].

Studies under the World Englishes paradigm mostly accounts for English production (Pillai, 2014; Pillai & Salaemae, 2012) or intelligibility (Deterding & Kirkpatrick, 2006; Ike, 2012). Thus, for the relation between perception and production of English vowels, this study borrows the models from Second Language Acquisition theories since they are more established at explaining the connection between perception and production. The major models used to explain the perception and production were advocated by SLA researchers such as Flege (1995), Best and Tyler (2007) and Escudero (2005). Since these models are rooted in SLA theories and where native-like proficiency is considered as ultimate attainment (Flege, 1995), comparison to native English varieties are common for studies using these models (Perwitasari et al., 2016; Perwitasari et al., 2015). Since the current study is placed under the World Englishes paradigm, in which achieving native like pronunciation is not desired and where unique variety is appreciated, in describing the relationship between perception and production, comparison to native speakers is avoided. However, the audio stimuli used to collect the perception data was compared with the production of English vowels by Acehnese speakers (See section 6.2.1). This comparison was not made to articulate the difference between the vowel quality produced by the American English speaker and Acehnese Indonesian speakers, but to determine which variable influenced the perception of English vowel contrasts.

2.7 Conclusion

This section has presented the literature reviews underlying the framework used in the current study. Three models under Global English (Three Circles, Dynamic English and English as a Lingua Franca) have been evaluated and compared with Kachru's Three Circles model. The World Englishes paradigm was selected as the model to explain the English variety spoken by students in Islamic boarding schools. World Englishes paradigm was selected because it fits the context of the speakers in the current study. Emergent English varieties in countries adjacent to Indonesia, such as Malaysia, Brunei and Singapore have also been presented to give an overview of other established emergent varieties in the region. Previous studies on the production of English vowels by Indonesian speakers (Javanese, Sundanese and Acehnese) were also presented. These studies were reviewed to see how similarly and differently Indonesian with various local language produced English vowels. The gaps in these studies were highlighted which the current study attempts to fill. In addition to studies on the production of English vowels, studies investigating perception of English vowel contrast was also discussed. Three models under Second Language Acquisition were reviewed and selected to explain the relationship between production and perception, Speech Learning Model, Perceptual Assimilation Model and Second Language Linguistic Perception. The next chapter provides the explanation of the method used to collect and analyse the data for this study.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This section begins by introducing the research paradigm within which the present study is situated, followed by an explanation of the selection criteria for the participants. The procedure of the materials used to elicit the data and the methods of data collection and analysis are then presented. This chapter also provides details of the ethical considerations involved in this study.

3.2 Research Design

This study is situated under the positivist paradigm. Under this paradigm, ontologically, an external reality exists and can be discovered and understood (Antonina, 2017; Kaboub, 2008). The production and perception of sounds are viewed as an existing external reality which can be observed and understood. The epistemological position of this paradigm is that the relationship between reality and the observer is distant in that the observer does not have any influence on the observed object (Grix, 2004). Thus, the researcher did not attempt to influence the findings of the participants' production and perception of sounds when collecting the production and perception data. In this regard, the researcher only acted as an observer who conducted an experiment on the observed object.

Axiologically, the nature of this study was value-free, excluding the author's own values when conducting the research to maintain an objective stance. Data under this paradigm were collected using scientific methods and further analysed based on a statistical analysis to achieve objectivity and neutrality during the research process (Grix, 2004). The scientific method employed in this study was designed in order to avoid any bias and ensure objectivity during data collection. Methodologically, according to this paradigm, the truth can only be uncovered by testing the hypothesis through scientific experiments

and manipulation of the observed condition. The implication of using the positivist paradigm was the use of a quantitative approach and the data was analysed statistically.

3.3 Participants

As explained in Chapter 1, this study examines the students at a Modern Islamic Boarding (MIB) school. At these schools, students are more likely to be fluent in Acehnese and Indonesian but have to communicate in English. The respondents in this study are students at one of the top five MIB schools in Banda Aceh, the capital city of Aceh Province (Fokusaceh, 2020). The students from this school also come from various towns and districts in Aceh.

The sample was selected using the purposive sampling technique. Using this technique, members of the population were selected based on a specific criteria (Campbell et al., 2020). In this study, only second- and third-year high school students who were fluent in English, Acehnese, and Indonesian and came from the west coast of Aceh were selected.

Initially, second and third-year students were selected to participate in this study because they would have developed a better aptitude in English communication skills, compared to the first-year students who would just be beginning to use English in their daily conversations. Subsequently, fluency in Indonesian, Acehnese and English was included as another criterion because this study looks at the possible influence of Indonesian and Acehnese vowels on their production of English vowels. It would be difficult to ascertain the influence of Indonesian and Acehnese on English if the participants were not fluent in these three languages. To determine fluency in English, only students with an English school report score of 70 to 100 marks in the last two semesters were considered for selection. Furthermore, Acehnese fluency was determined by distributing questionnaires confirming their language use in various settings (see Appendix A). The questionnaires attempted to find out if the students spoke Acehnese with their parents, siblings, relatives,

and friends. If they indicated that they used Acehnese in most of these settings, I confirmed their fluency by speaking Acehnese with them. In relation to Indonesian, no special test was administered to determine their fluency because Bahasa Indonesia is the medium of instruction in the school. Students attending the school need to be fluent in Bahasa Indonesia in order to communicate with and learn from school teachers.

Finally, this study specifically selected students from the west coast of Aceh (Aceh Jaya, Aceh Barat, and Nagan Raya) because the schools in this region have a significant number of students from these areas. Considering the fact that there are different Acehnese accents (Durie, 1985), and phonetic research is sensitive to accent, the study focused on one specific Acehnese accent. In addition, the Aceh Utara accent has been previously studied (Asyik, 1987; Durie, 1985; Yusuf, 2013), while the west coast Acehnese is relatively under-researched (Tanzir Masykar, Roni Agusmaniza, et al., 2021; Zulfadhli, 2015). Thus, based on these criteria, a total of 32 students (16 males and 16 females) were selected. However, the data from three students (two females and one male) were excluded from the analysis because they did not produce several of the target English monophthongs. Thus, only data from 29 students (14 females and 15 males) were eventually analysed.

3.4 Data Collection

This section describes the materials that were used to elicit the data for the production and perception tasks. Two types of data were collected using the production and perception tasks respectively. The production tasks elicited the data which were used to examine vowel quality and length in English, Indonesian, and Acehnese. The perception tasks were used to examine the extent to which the students were able to distinguish English vowel pairs. For the production data, vowels produced in three languages,

English, Acehnese and Indonesian, were recorded, while for the perception data, only students' perception of English vowels was examined.

3.4.1 Production Test

The production test was administered before the perception test to avoid students getting an idea on how to pronounce the English words during the perception test. Three tasks were administered in the production task for each language in the following order, English, Indonesian, and Acehnese. All students took the task one at a time to avoid the possibility of peer influence. The production task was conducted on three consecutive days. Three research assistants helped during the data collection period.

All students went through two phases, the familiarisation phase, and the recording phase. Both phases were conducted in two different rooms in the same building at the school. In the familiarisation phase, students were asked to familiarise themselves with the target words in a separate room using one laptop per student at a time. A research assistant guided them on what to do in the actual production task. The students were asked to read the sentences at a normal speed. They could practise reading all the words in the list, but they were not allowed to ask how to pronounce the target words. The procedure for the Acehnese production task was also introduced at this phase. After they were familiar with the target words and the procedure for recording, they moved to the recording room.

At the recording stage, the recording started with the name of the student, the language, and the recording number. This was to enable a systematic classification of the data for analysis. The production task was recorded with a Zoom H6 recorder connected to a headset microphone. The recordings were sampled at 44,000 Hz and 16-bit rate and recorded in the .WAV format to maintain the quality of the recordings. The microphone was placed comfortably in front of their mouth about 5 cm away from the lips. The close

position of the microphone to the mouth allowed the recorder to capture clearer recordings with minimal ambient noise.

The students were required to read target words in a carrier sentence for each respective language:

I say___again (for English)

Saya bilang___lagi [I say___again] (for Indonesian)

*Nyoe lam Bahasa Aceh*___ [This in Acehnese is___] (For Acehnese).

The target words in the carrier sentences were presented on a computer screen using PowerPoint slides, except for Acehnese. For the latter, images associated with the target words and cue words in Indonesian were presented on the slides. Picture and cue words were provided for the Acehnese production task because unlike Indonesian, Acehnese does not have a standard written form. This was done to minimise the possibility of the students treating written target Acehnese words as they might do in Indonesian. Due to the nature of Acehnese language and in order to elicit more naturally produced words containing the target vowels, The target Acehnese words were also placed at the end of the sentence. During the recording stage, the student moved the slides containing the images at their own pace.

During the recording session, the students were recorded reading each word three times. This was done for all languages, starting from English, Indonesian, and Acehnese. The students were asked to pause for about three seconds between each sentence containing the target word. They finished all the words in one take before proceeding to the second and then third repetitions. There was a short break of about one minute between the first, second and third take and a short break of up to five minutes between each language

production task. The recordings were conducted in the language soundproofed laboratory at the school.

Elicitation materials

Three types of materials were administered in this study to elicit the production of English, Indonesian, and Acehnese. Firstly, the elicitation materials for the English production task contained ten words targeting English monophthongs. The monophthongs used in the study were based on American English monophthongs as described in Ladefoged (2003). The Indonesian school curriculum for English does not specify which English variety should be used (Lauder, 2010). Lauder (2020) also argued that there is no single standard model of English variety for Indonesian learners to adopt which can be either American English, British English or both. However, this study opted for American English (see 2.3.2.4) due to its dominant position compared to British English, especially outside the United Kingdom (Gonçalves et al., 2018). In Indonesia, major TV stations also tend to air American movies and as such Indonesian speakers are more exposed to American English than British English.

For the target English words, only words in a CVC (consonant vowel consonant) context were selected. The initial and final consonants were plosives or fricatives so as not to obstruct the quality and length of the following target vowels. The plosive consonants included bilabial dentals, alveolars and velar plosives, [p], [b], [t], [d], [k], and [g], while the fricatives were [f], [v], and [h]. The target vowels and words used for the English production test are presented in Table. 3.1

Table 3.1: Word list for English monophthongs

No	Target Words	Target Vowel	Lexical Set
1	pot	[ɑ:]	LOT
2	bit	[ɪ]	KIT
3	beat	[i:]	FLEECE
4	bet	[ɛ]	DRESS

5	bat	[æ]	TRAP
6	bud	[ʌ]	STRUT
7	foot	[ʊ]	FOOT
8	food	[uː]	GOOSE
9	port	[ɔː]	FORCE
10	bird	[ɝ]	NURSE

Secondly, similar to English, the target monophthongs for the Acehnese production task were also placed in a CVC context. Eleven words were selected, targeting eleven Acehnese oral monophthongs described in previous studies (Asyik, 1987; Durie, 1985; Pillai & Yusuf, 2012). The words were adopted from the study by Pillai and Yusuf (2012) based on the northern Acehnese dialect (see section 2.3.2.2). The target vowels and words used for the Acehnese production task are shown in Table. 3.2.

Table 3.2: Word list for Acehnese monophthongs

No	Target Words	Word Class	Target Vowel	English
1	*cit*	adverb	[i]	'too, also'
2	*cut*	noun	[u]	'small, title for women of noble descent'
3	*pét*	verb	[e]	'close / shut eyes'
4	*tet*	verb	[ə]	'burn'
5	*pôt*	verb	[o]	'blow, to fan'
6	*göt*	adjective	[ʌ]	'good, fine'
7	*cop*	verb	[ɔ]	'sew'
8	*pat*	conjunction	[a]	'where'
9	*peut*	adjective	[ɯ]	'four'
10	*cèt*	verb	[ɛ]	'paint'

The third data elicitation was the Indonesian production task. The task contained seven Indonesian words targeting six Indonesian monophthongs described in previous studies (Soderberg & Olson, 2008; Zanten & Heuven, 1984). Since previous studies on Indonesian vowels did not include the [ɔ] vowel, the word *bobok* [sleep] was included in

the list to elicit this *vowel*. Most Indonesian words are disyllabic, and thus, all words in the Indonesian production task consisted of two-syllable words similar to what was used in previous studies (Soderberg & Olson, 2008; Zanten, 1986; Zanten & Heuven, 1984). Thus, the target monophthongs were placed in CVCVC, with target vowels placed in the last CVC. The target vowels and words used for the Indonesian production task can be observed in Table. 3.3

Table 3.3: Word list for Indonesian monophthongs

No	Target Words	Target Vowel	Glossary
1	*bib**i**t*	[i]	'seed'
2	*bub**u**k*	[u]	'powder'
3	*beb**e**k*	[ɛ]	'duck'
4	*bab**a**t*	[a]	'food from animal abdomen'
5	*bob**o**t*	[ɔ]	'weight'
6	*bob**o**k*	[o]	'sleep'
7	*tet**a**p*	[ə]	'remain'

3.4.2 Perception Task

Five English vowel pairs were used for the perception task: [ɪ] – [iː] (bit - beat), [ɛ] – [æ] (be - bag), [ʊ] – [uː] (foot - food), [ʌ] – [ɑː] (bud – bard), and [ɑː] – [ɔː] (pot - port). Each pair is referred in this thesis as VP1 for [ɪ] – [iː], VP2 for [ɛ] – [æ], VP3 for [ʌ] – [ɑː], VP4 for [ʊ] – [uː] and VP5 for [ɑː] – [ɔː]. The words and vowel pairs used are presented in Table 3.4. An additional four words for the [ɪ] – [iː] (fit - feet) and [ɛ] – [æ] (bet - back) pairs were included for students to practice and become familiar with the task. The recordings for the three perception tasks were provided by a 26-year-old male American speaker. The recorded audio containing target words were used for the three perception tasks, AX, ABX and FCI. The recordings were carried out in one of the rooms at the Syiah Kuala University Language Centre. The speaker, who consented to be recorded, was born and raised in Oregon and self- reported to having a Western American accent.

The speaker was working as an English teacher at Syiah Kuala University at the time of the recordings and had been in Aceh for one year.

Table 3.3: Vowel pairs and words for perception task

		Vowel Pairs	Word Pairs
Practice Trial	VP1	ɪ - iː	*fit - feet*
	VP2	e - æ	*bet - back*
Actual Trial	VP1	ɪ - iː	*bit - beat*
	VP2	e - æ	*beg - bag*
	VP3	ʊ - uː	*foot - food*
	VP4	ʌ - ɑː	*bud - bard*
	VP5	ɑː - ɔː	*pot - port*

The perception task was conducted after all the students finished recording the production task. The perception task consists of three cycles, one Forced Choice Identification (FCI) task and two discrimination tasks (ABX & AX). The order in which tasks were conducted was ABX, AX and FCI. Iverson and Evans (2009) suggest that the identification task should always be placed last because, in the identification task, students are presented with the words used in the previous two tasks. This order also minimises their perception from being influenced by their lexical knowledge (see Alispahic et al. (2017).

The students completed the perception task as a pen and paper task with the audio stimuli played through a computer connected to a loudspeaker and placed in front of the room. All students completed the task together in one of the rooms in the school language laboratory. The laboratory was a quiet room, could accommodate all students at once, and was considered suitable for playing the audio. The volume was set at a comfortable level and adjusted to accommodate all the students in the room. This is similar to the setup by Mora (2005). The answer sheets for three tasks were combined into a single booklet with

a cover for students to write their names and student number (see Appendix X for answer sheet).

Before the task began, students were told that they would take three different task with a break of five minutes in between each task. In each task, the students were given a familiarisation set of stimuli before taking the actual task. The data from the familiarisation task were excluded from this study. Students were also informed that they would listen to several trials, and, in each trial, they would listen to English words played in a sequence at a specific interval. To encourage the students to do their best in the perception task, they were told that the first three students who achieved high scores would be rewarded with gifts. A reward is important in education to stimulate motivation and seriousness when doing a task. Van Hessen and Schouten (1999) also rewarded their participants based on the number of correct answers they got. The next three sections describe sthe materials and specific procedures of the three tasks.

ABX Task

For the ABX task, the students were asked to decide whether the third word (X) they listened to in the sequence was similar to the first (A) or second word (B) which they had previously heard. Within five seconds after they heard the last word, they had to make their decision and tick the first or second box in their answer sheet while getting ready for the next trial. If students were unable to make their decision, they could guess, and they were told that no answer should be left blank. They could always ask the research assistant questions regarding the procedures before the actual task was played.

As for the design of the ABX task, the 14 recorded words provided by the American speakers were sequenced into four possible orders for each vowel pair, ABA, ABB, BAA, and BAB. For example, for the vowel pairs [ɪ] – [iː], using the word *'fit'* and *'feet'*, the four possible sequences would be, *'fit - feet - fit', 'fit - feet – feet', 'feet - fit – fit',* and

'feet - fit – feet'. An Inter Stimulus Interval (ISI) of 1 second was inserted between the words, and an Inter Trial Interval (ITI) of three seconds separated each trial. A popular open-source Audacity software (Audacity Team, 2021) was used to combine the words into a single trial. Each trial was duplicated into five repetitions producing a total of 20 trials for each vowel pair. These 20 trials were randomly placed in a continuum with the ITI of five seconds to give the students time to tick the provided box in the answer sheet. The 20 trials were combined into one file and saved into WAV format. The combined 20 randomised trials of each vowel pair were referred to as a block. There were two blocks for the practice materials and five blocks for actual materials. The word sequences designed for the practice and actual materials are shown in Table 3.5 and Table 3.6

Table 3.4: ABX design for practice materials

Block	Vowel Pair	Order	Stimuli	Repetition	Total
I	[ɪ] – [iː]	ABA	fit - feet - fit	5 rep.	20
		ABB	fit - feet - feet	5 rep.	
		BAA	feet - fit - fit	5 rep.	
		BAB	feet - fit - feet	5 rep.	
II	[e] – [æ]	ABA	bet - back - bet	5 rep.	20
		ABB	bet - back - back	5 rep.	
		BAA	back - bet - bet	5 rep.	
		BAB	back - bet - back	5 rep.	
Total					40

Table 3.5: ABX design for actual materials

Block	Vowel Pair	Order	Stimuli	Repetition	Total
III	[ɪ] - [iː]	ABA	bit - beat - bit	5 rep.	20
		ABB	bit - beat - beat	5 rep.	
		BAA	beat - bit - bit	5 rep.	
		BAB	beat - bit beat	5 rep.	
IV	[ɛ] - [æ]	ABA	beg - bag - beg	5 rep.	20
		ABB	beg - bag - bag	5 rep.	
		BAA	bag - beg - beg	5 rep.	
		BAB	bag - beg - beg	5 rep.	

V	[ʌ] – [ɑː]	ABA	bud - bard - bud	5 rep.	20
		ABB	bud - bard - bard	5 rep.	
		BAA	bard - bud - bud	5 rep.	
		BAB	bard - bud - bard	5 rep.	
VI	[ʊ] – [uː]	ABA	foot - food - foot	5 rep.	20
		ABB	foot - food - food	5 rep.	
		BAA	food - foot - foot	5 rep.	
		BAB	food - foot - food	5 rep.	
VII	[ɒ] – [ɔː]	ABA	pot - port - pot	5 rep.	20
		ABB	pot - port - port	5 rep.	
		BAA	port - pot - pot	5 rep.	
		BAB	port - pot - port	5 rep.	
Total					100

AX Task

After they finished the ABX task, the students went on to the next page and completed the AX task. As the design of the AX task was different from the ABX task, new instructions were given prior to the task. They were told that in this task, two words were sequenced in a trial. Each trial consisted of the same or different words. Students were asked to decide whether the word pairs were the same or different. After they made their decision, they had to mark their answer in the 'same' or 'different' box provided in the answer sheet. Just like in the ABX task, the students also got the chance to familiarise themselves with the task format. They did the practice task before taking the actual task and marked the answers in two separate answer sheets provided in the booklet. They listened to a total of 40 trials in the practice task and 100 trials in the actual task. The trials were divided into a block with 20 trials in each block.

As for the design of the ABX task, the same 14 English words produced by the native American speaker were used in the ABX task. In the AX task, the word X could be the same or different from the A word. Four possible alternatives were sequenced in a single trial, AA, AB, BA, BB. For example, for the VP2 [ɪ] – [iː], the word pairs used were 'fit' and 'feet' and the four possible sequences would be, '*fit - fit', 'fit – feet', 'feet – fit',* and

'feet - feet'. A one-second Inter Stimulus Interval (ISI) was inserted in between the two stimuli, while a five second Inter Trial Interval (ITI) was inserted between the trials. The stimuli were combined using the Audacity software (Audacity Team, 2021).

Like the ABX task, the trials for the AX task consisted of the practice task and the actual task. The practice task contained 40 trials divided into two blocks. Each block comprised 20 trials with five repetitions for each possible sequence. The actual task contained 100 trials. The trials were divided into five blocks containing 20 trials of vowel pairs in each block. The design sequence of the practice and actual material can be observed in Table 3.7 and Table 3.8

Table 3.6: AX design for practice materials

Block	Vowel Pair	Order	Stimuli	Repetition	Total
I	[ɪ] – [iː]	AA	fit - fit	5 rep.	20
		AB	fit - feet	5 rep.	
		BA	feet - fit	5 rep.	
		BB	feet - feet	5 rep.	
II	[ɛ] – [æ]	AA	bet -bet	5 rep.	20
		AB	bet - back	5 rep.	
		BA	back - bet	5 rep.	
		BB	bet - bet	5 rep.	
Total					40

Table 3.7: AX design for actual materials

Block	Vowel Pair	Order	Stimuli	Rep.	Total
III	[ɪ] – [iː]	AA	bit - bit	5 rep.	20
		AB	bit - beat	5 rep.	
		BA	beat - bit	5 rep.	
		BB	beat - beat	5 rep.	
IV	[ɛ] – [æ]	AA	beg - beg	5 rep.	20
		AB	beg - bag	5 rep.	
		BA	bag - beg	5 rep.	
		BB	bag - bag	5 rep.	
V	[ʌ] – [ɑː]	AA	bud - bud	5 rep.	20
		AB	bud - bard	5 rep.	
		BA	bard - bud	5 rep.	

		BB	bard - bard	5 rep.	
VI	[ʊ] – [uː]	AA	foot - foot	5 rep.	20
		AB	foot - food	5 rep.	
		BA	food - foot	5 rep.	
		BB	food - food	5 rep.	
VII	[ɑː] - [ɔː]	AA	pot - pot	5 rep.	20
		AB	pot - port	5 rep.	
		BA	port - pot	5 rep.	

		BB	port – port	5 rep.	
Total					100

FCI Task

The identification task was the last task the students completed in the perception task. In this task, the students were told that they had to listen to one word, and the word was repeated three times. After that, they had to identify and mark which of the two possible words on their answer sheet corresponded to the word they had heard in the trial. A total of five blocks were played in the actual task and two blocks in the practice task. Each block consisted of 12 trials.

The task utilised the same 14 words from the previous task, four for the practice task and 10 for the actual task. For example, a student listens to a trial that contains the word 'bit' [bɪt] three times and then they ticked (√) in the answer sheet if the student think that was the word in the recording. Five copies of the trial 'bit' and the trial 'beat' were generated using the Audacity (Audacity Team, 2021), producing a total of 10 trials. The ten trials were randomly presented into a single block. Like the previous task, there were seven blocks of vowel pairs in the forced choice identification task. The students listened to two blocks in the practice task and five blocks in the actual task. The words and design of the practice task are shown in Table 1.9, while the actual task is presented in Table 1.10.

Table 3.8: FCI design for practice materials

Block	Stimuli Heard	Target Vowel	Repetition	Choices
I	fit	[ɪ]	6	fit
				feet

	feet	[iː]	6	fit
				feet
II	bet	[ɛ]	6 rep.	bet
				back
	back	[æ]	6	bet
				back
Total			24	

Table 3.9: FCI design for actual materials

Block	Stimuli Heard	Target Vowel	Repetition	Choices
III	bit	[ɪ]	6	bit
				beat
	beat	[iː]	6	bit
				beat
IV	beg	[ɛ]	6	beg
				bag
	bag	[æ]	6	beg
				bag
V	bud	[ʌ]	6	bud
				bard
	bard	[ɑː]	6	bud
				bard
VI	foot	[ʊ]	6	foot
				food
	food	[uː]	6	foot
				food
VII	pot	[ɒ]	6	pot
				port
	port	[ɔː]	6	pot
				port
Total			60	

3.5 Ethical Considerations

A letter from Universiti Malaya about the need to carry out the data collection as part of a PhD study was obtained and used to obtain permissions from the local ministry of

religious affairs in Banda Aceh. Islamic schools in Indonesia are under the Ministry of Religious Affairs. This ministry then issued a letter to the school headmaster to grant permission for the researcher to conduct the study at the school. The headmaster assigned a male dormitory teacher and a female supervisor teacher to assist throughout the study. The school also allowed the researcher to use some rooms in the school laboratory building, which was quiet and deemed adequate for conducting the task.

Throughout the data collection, the teachers accompanied the researcher and helped with the permission needed from the dormitory teachers. In boarding schools, dormitory teachers act like parents and parents entrust their children to the teachers (*ustadz)* as their guardians. Thus, permission to collect the data from the students was obtained from the dormitory teachers. All rules imposed at the school were also upheld throughout the study. These included separating boys and girls between the aisles during the perception task, having a female assistant when conducting production tasks with the female students, having breaks for prayers and lunch break, and adhering to the school language policy when to use English when communicating with the students. This was with the exception of using Indonesian when explaining the technicalities required (which was approved by the school supervisor). Written consent was obtained from the student participants prior to administering the tasks. Written consent was also obtained from the American speaker who provided the audio for the perception task.

3.6 Data Analysis

The following section explains how the data from the production and perception tasks were analysed. The analysis of the production data from the three languages, which addressed the first (RQ1), second (RQ2) and third research questions (RQ3), involved four stages: file naming, isolation and annotation, acoustic measurements, conversion of the formant values, and statistical analysis.

In the production data analysis, the first stage was file naming of the recorded files. All production data was saved in .WAV format with a total recording time of between 90 and 120 seconds for each set. There were three sets of data produced by 29 students making a total of 87 sets for each production task, for English, Indonesian, and Acehnese. All recorded WAV files of each speaker were given names: student name, language, and take number. The files were then saved into folders based on the speaker's name folders. These naming systems would later be helpful when measuring the monophthongs, generating formants and durational data, and revisiting the data whenever outliers and peculiar data were found during data sorting, vowel space projection, and statistical analysis.

The second stage was vowel isolation and annotation. After all the recorded files had been labelled, each set was analysed in Praat version 6.2.14 (Boersma & Weenink, 2022) to obtain formant and durational measurements. Prior to measurement, the data were first annotated in Praat. Using Praat, the target vowels were isolated by determining the onset and offset of each target vowel. The target sound within the boundary was then annotated with each respective word and target vowel using the phonetic symbol. This process was done manually for each target vowel within the set. Each file was saved as TextGrid file and named after the corresponding audio file.

The third stage was the formant and durational measurement. Once vowel isolation and annotation for all the words containing the target vowels for each language were completed, the data was ready for formant and durational measurement. Only the first (F1) and second formants (F2) were measured to determine the quality of vowels. These two formants could be used to plot the quality of the vowels into the F2 – F1 vowel space. Even though the quality of a vowel can be represented by many formants, the two first formants are enough to represent the vowel quality in the vowel space. If F1 is similar in

between two vowels, they must differ in F2 to indicate their differences (Be u , 2021). The onset and offset of the vowel were determined by manually inspecting the waveform and spectrogram of the target vowel. Figure 3.1. illustrates the onset and offset of the vowel [æ] in *bat*. The onset of the [æ] vowel starts shortly after the plosion in the bilabial plosive [b] and ends shortly before the next plosion in the alveolar plosive [t]. The onset of the vowel is indicated by the first marking line while the offset is shown by the last marking line as seen in Figure 3.1. These marking lines were later used to retrieve the formants and durational data automatically using a script.

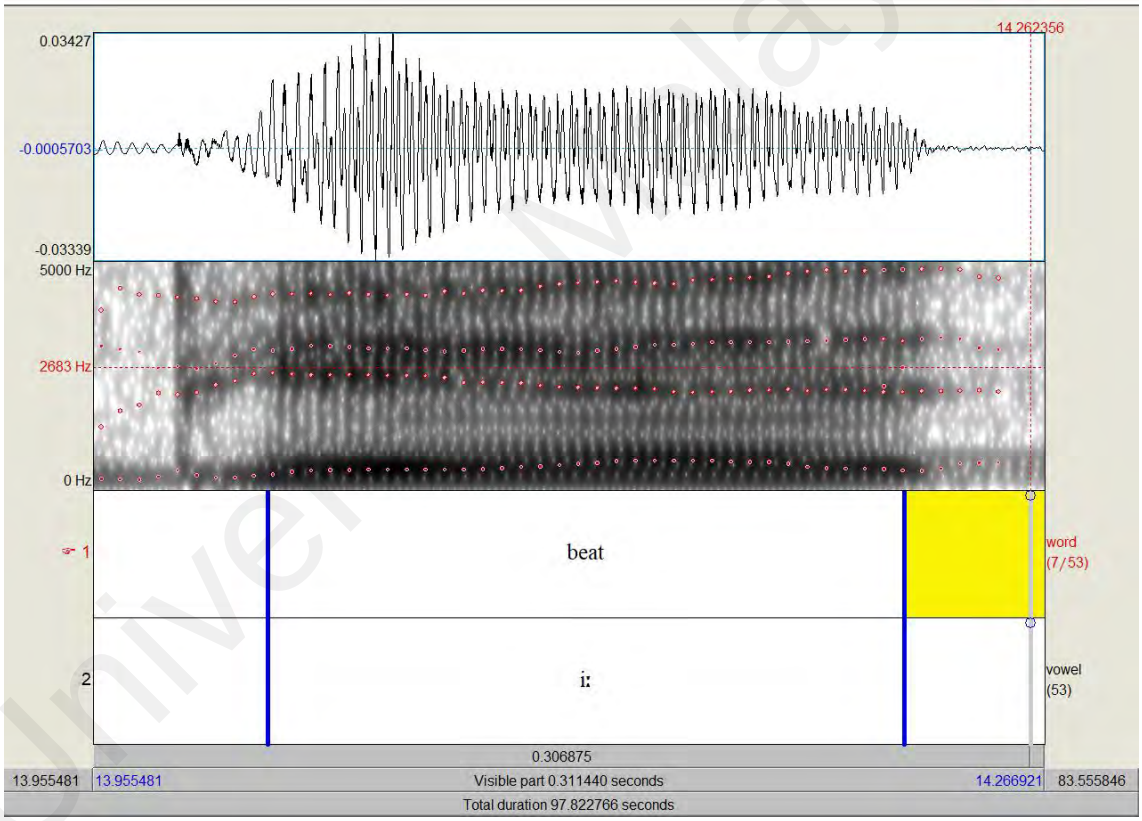

Figure 3.1: Onset and offset of English [iː] in beat

The measurement of F1, F2 and duration was done automatically using the script developed by Buss (2014) retrieved from his github repository (Buss, 2014). The script employed the "*to formant (burg)"* method that considers the physiological difference

between males and females (5500 Hz predefined for females, 5000 for males). The script harnessed the data based on the word and vowel annotated in the TextGrid. The script measured the first three formants at the start, mid, and end point of the vowel and attached the results to a text file. Average formant values for each vowel were automatically calculated and appended to a formant file. The text file was automatically placed in the same location of the text-grid directory and was named correspondingly. In addition to formants, the script also collected the temporal data for each vowel within the set. Since the script only collected the target vowels' onset, mid and offset times, vowel duration for each vowel was manually generated by subtracting the final and initial time in milliseconds in Excel.

Prior to converting the formant values and performing statistical analysis, the data were tabulated, organised and named accordingly in Excel. Each formant value of each target vowel was also first averaged for each speaker. Bark scale is used because it better represents how humans hear the vowels. The bark scale "is thought to be a good approximation of the actual frequency analysis performed by the ear" (Kent & Read, 2002, p. 113). The scale also represents formant values in a smaller unit compared to hundreds and thousands in Hertz. Thus, the distance between each vowel is easier to conceive. To convert average F1 and F2 Hertz values to Bark, the following formula from Zwicker and Terhardt (1980, p. 1524) was applied to the formant data in the excel table:

$$\mathbf{Zc = 13*ATAN\ (0.00076*F1) + 3.5*ATAN\ ((F1/7500) * (F1/7500))}$$

Where:

- Zc is critical-band rate in Bark
- ATAN is applied to numbers in radians
- F is frequency in kHz

After F1 and F2 values were converted to Bark, the data analysis proceeded to vowel space plotting and statistical analysis. Euclidean distance (ED) was also calculated to determine the distance of all vowels from the centroid. Following Deterding (1997) the average ED was calculated by averaging ED values of all vowels except the central vowel [ɜ˞]. ED is used to determine how peripheral the English vowel space is compared to previous studies and to Acehnese and Indonesian vowel in the current study. To answer the RQ1, the average formant value for each target monophthongs in the three languages were plotted against F2–F1 values. This plot represents the quality and location of each monophthong within the vowel space of each respective language. The durational data for English vowels were presented in a line graph. As for the statistical analysis, the study employs three primary statistical analyses to answer different research questions—Independent sample t-test, paired sample t-test, analysis of variance (ANNOVA), and Bonferroni Post Hoc test. The first statistical analysis conducted was the paired sample t-test. This test was selected to answer the RQ2, which sought to learn whether the students distinguish typical English vowel pairs in terms of quality and length. For quality contrast, the statistical analysis compared the mean frequencies of F1 and F2 for the vowel pairs across 29 students. For the length contrast, the statistical analysis compared the mean duration of each vowel pair produced by the students.

ANOVA was used to answer RQ3; it was used to ascertain whether the production of English vowels equivalent to Acehnese and Indonesian vowels exhibited any significant differences. The mean F1 and F2 of each monophthong in the three languages were compared using ANOVA. The findings helped determine which English vowel is closer to their first language (Acehnese) and second language (Indonesian). A post hoc comparison using a Bonferrroni adjustment was also conducted where statistically significant findings were produced. Effect size was also calculated to determine the practical significance of a research outcome.

The analysis of English vowel perception involved data tabulation, charting and statistical analysis. This analysis was carried out to answer the fourth and fifth research questions (RQ4 and RQ5). The first stage in the perception data analysis involved tabulating the perception data into Excel. Since there were three perception data from three perception tasks, the data were tabulated one by one. For the perception task, a score of 1 was given each time the student had the correct perception and 0 if otherwise. For the ABX and AX task, the highest score the student could get for each pair tested was 20 points, while for FCI, it was 12 points, representing the number of total tokens played for each pair. The average scores for each pair in the three-perception task were calculated by calculating the mean score obtained by all the students.

The second stage was charting the data. To produce the chart, the average scores for each vowel pair were projected into the bar chart and compared against the three tasks to observe their performance in the perception task. These charts were produced to depict if the students perceived the five English vowel pairs differently across three types of tasks (RQ4). To determine if the perception of each pair was significantly different from one another, a repeated measure ANOVA was conducted across each perception task. A post hoc comparison using a Bonferrroni adjustment was conducted should statistically significant analysis be produced. The means value for minimal pairs and repetition for each vowel construct were also compared. The post hoc test was conducted following this to see which order of presentation is perceived as significantly different from the others should the ANOVA test indicate any significant findings.

In order to answer RQ5, the extent to which the production of their English vowels related to the perception of the same vowels, the study used the descriptive analysis. The score in the perception task was compared against the results from the production task. First, the result of statistical analysis of English vowel contrast was classified into PAM

classification together with its prediction. Second, this classification together with its prediction is compared against the results from perception test on the three perception tests. This analysis examined whether students' ability to distinguish English vowel pairs in the perception task reflected their ability to do so in the production task. Additionally, the result from the perception test was. Also compared to the American speakers providing the recording for perception test to see if if the perception of vowel contrast was bounded by the cue provided by the American speaker or by the production of the English vowel contrasts.

3.7 Vocabulary Size Test

Before the speakers participated in the production and perception task, they sat for a vocabulary size test to determine their level of proficiency and to divide their performance into groups. Following the vocabulary size test, their familiarity with the words used in the test was also reviewed. The vocabulary test was developed by Meara (1992) and has been used by previous researchers to determine speakers' proficiency. Though it does not account for their speech proficiency, the test provides a rough picture of how each student performs in comparison to other speakers in the group. This vocabulary test was also used by researchers working on a similar study (Perwitasari, 2018, 2019). The list of words in the vocabulary test contains English words and non-words. Speakers scored 100 if they were able to identify all 40 English words presented in the test. Each time the student missed ticking any English word, their scores would be subtracted by -2.5 and a -5 is subtracted if they had ticked the non-words. The overall score was calculated based on the number of correct answers. Speakers' classification based on the vocabulary test was used to interpret the perception data in AX, ABX and FCI.

Based on the vocabulary size test findings, the speakers were divided into three groups: high, mid, and low performing speakers. The comparison can be observed in Table 3.10.

Each group had a score difference of 20 compared to the group above its level. The groups were divided categorically based on the score of all students. Speakers with a score between 80 to 100 were in the high-performing group, 60 to 79 were in the mid performing group, and those who scored below 60 were considered to be the low-performing group. The high-performing speakers had an average vocabulary size test of 85.2 with 95 as the highest score and 80 as the lowest score. The average score for the medium-performing speakers was about 70, with 75 and 60 as the highest and lowest score, respectively. The low-performing speakers had an average score of 42, with the highest score at 55 and the lowest score at 30. The findings from the vocabulary size test would be used to observe how each group of speakers perform in the perception and production test and how their proficiency is related to their production and perception relationship.

Table 3.10: Classification of speakers into high, mid and low group based on vocabulary size test

	Number of Speakers	Average Score	Highest Score	Lowest Score
High-performing	10	85.2	95	80
Medium-performing	11	66.9	75	60
Low-performing	8	41.9	55	30

The detailed distribution of speakers' vocabulary size tasks can be observed in Table 3.11. It is clear from the table that most speakers (a total of 72.4%) are in the high and medium category. A lesser number of speakers, i.e., only 27.6% were in the low performing category. Most speakers in the high-performing category scored between 80 and 88, while in the medium-performing category, the majority scored between 60 and 67.5 Only three speakers scored above 90 in the high category, and three speakers above 70 in the medium-performing category. The results for the lower category were more spread out, ranging from 30 to 55. Three speakers were placed in the lower end of the table with a

score of 30 to 32.5, and three speakers at the upper end of the table with a 52.5 – 55 score. Only two speakers in the lower category score in the 40s range. These findings were expected because the speakers are in the second and third-year of their high school. They have been using English for two to three years during their stay at the boarding school. In addition, their scores in English language subject were above 70 in the previous two semesters.

Table 3.11: Detailed distribution of vocabulary size score into high, mid and low group

No	Higher Score		Medium Score		Lower Score	
	Code	%	Code	%	Code	%
1	FS1	95	FS3	75	FS16	55
2	FS6	90	MS1	72.5	FS10	52.5
3	MS8	90	MS7	70	FS11	52.5
4	MS6	87.5	FS14	67.5	FS9	42.5
5	FS7	85	MS3	67.5	MS14	40
6	MS2	85	FS15	65	MS4	32.5
7	FS4	82.5	FS13	62.5	MS10	30
8	FS5	82.5	MS5	62.5	MS11	30
9	FS2	80	MS9	62.5		
10	FS12	80	MS15	62.5		
11			MS13	60		

3.8 Word Familiarity Index

In addition to grouping their proficiency into three groups based on the vocabulary test, the speakers were also asked if they knew any of the words used in the perception test. The review was conducted after they participated in the perception test. The review required the speakers to put a checkmark on words they were familiar with. The findings can be observed in Table 3.12. For the word familiarity test, the speakers were also divided into high, medium, and low familiarity index categories. The division were also made categorically based on the score of all students. The majority of speakers (13) fell under on the low category compared to 9 and 5 in the medium and high category

respectively. The number of lower category speakers was higher than the combined total of speakers in the medium and high categories. The speakers in the higher group were familiar with 8.6 of the total ten words used in the test. The average word familiarity for medium and low groups was 6.4 and 4.3, respectively. The highest number of recognized words for each group was 10 for the high group, 7 for the medium group, and 5 for the low group. The lowest number of recognized words for each group was 8 for the high group, 5 for the medium group, and 2 for the low group.

Table 3.12: Word familiarity index for the high, medium and low group

		Average		Highest		Lowest	
	Total Speakers	%	Words	%	Word	%	Word
High	9	86	8.6	100	10	80	8
Medium	5	63	6.4	70	7	50	5
Low	13	43	4.3	50	5	20	2

The detailed distribution for the student familiarity test can be observed in Table 3.13. Two speakers in the highest group (FS14 & MS2) recognized all the words in the test, and only one student in the lowest group recognized only two words (MS11). Interestingly, most speakers (9 out of 15) in the lower category recognized 50% of the words used, while most speakers (6 out of 9) recognized 80% of the words in the higher category. The number of speakers in the medium category was almost equal. Three speakers recognized 70% of the words, and four speakers recognized 60% of the words.

Table 3.13: Detailed distribution of word familiarity index score for high, mid and low group

No	Higher Score			Medium Score			Lower Score		
	Code	%	Words	Code	%	Words	Code	%	Words
1	FS14	100	10	FS1	70	7	FS9	50	5
2	MS2	100	10	FS12	70	7	FS10	50	5
3	FS16	90	9	MS6	70	7	MS1	50	5
4	FS2	80	8	FS3	60	6	MS9	50	5
5	FS4	80	8	FS15	60	6	MS10	50	5

6	FS5	80	8	MS7	60	6	MS12	50	5
7	FS6	80	8	MS8	60	6	MS13	50	5
8	FS7	80	8				MS14	50	5
9	FS11	80	8				MS3	40	4
10							MS4	40	4
11							MS15	40	4
12							MS5	30	3
13							MS11	20	2

Table 3.14 provides information on the number and percentage of speakers who recognized the words used in the perception task. All speakers were familiar with the word *foot* [ʊ] and *food* [uː] while the words *bud* [ʌ] and *bard* [ɑː] are less recognised. The number of speakers familiar with the words *bud* [ʌ] and *bard* [ɑː] were only 9 and 6, respectively. It is quite surprising that four speakers claimed to be unfamiliar with the word *bag* [æ] considering the word would have been frequently used during their stay at the school. The number of speakers familiar with the other words in the list varies. Eighteen speakers reported being familiar with the word *beat* [iː] and *port* [ɔː], while 19 and 17 knew the words *pot* [ɑː] and *bit* [ɪ] respectively. The visual representation of the table from the most to the least familiar can be observed in Figure 3.2.

Table 3.14: Percentage and number of speakers familiar with words used in the study

Word	Vowel	# student	% student
foot	[ʊ]	29	100
food	[uː]	29	100
bag	[æ]	27	87
pot	[ɑː]	19	61
beat	[iː]	18	58
port	[ɔː]	18	58
bit	[ɪ]	17	55
beg	[e]	11	35
bud	[ʌ]	9	29
bard	[ɑː]	6	19

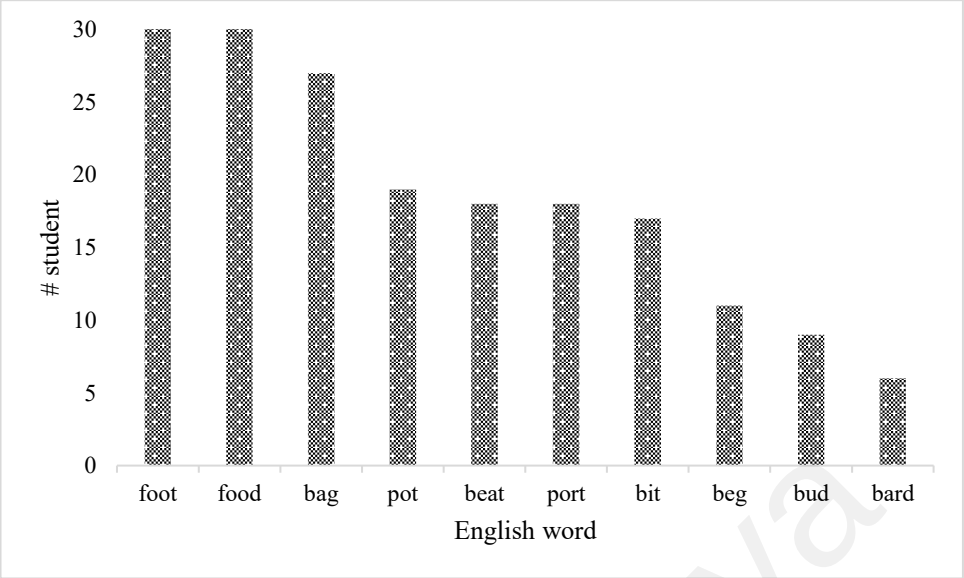

Figure 3.2: Number of speakers familiar with words used in the study

Table 3.15 presents the familiarity index based on the word pairs. The percentage of the speakers for each word is also included. Both words in Pair 3 (*foot – food*) are recognized equally at 100%. Both words in Pair 1 (*bit* - *beat*) and Pair 5 (*pot – port*) are equally recognized with only a difference of one student. The percentage of familiarity for both words in Pair 5 is 61% and 58%, while for Pair 2 is 55% and 58%. Pair 4 has the least familiarity index at 29% and 19% for both words in the task. The visual representation of the table for each pair can be observed in Figure 3.3.

Table 3.15: Word familiarity index by word pairs

	Word	Vowel	# student	% student
Pair 1	*bit*	ɪ	17	55
	beat	iː	18	58
Pair 2	*beg*	e	11	35
	bag	æ	27	87
Pair 3	*foot*	ʊ	31	100
	food	uː	31	100
Pair 4	*bud*	ʌ	9	29
	bard	ɑː	6	19
Pair 5	*pot*	ɑː	19	61
	port	ɔː	18	58

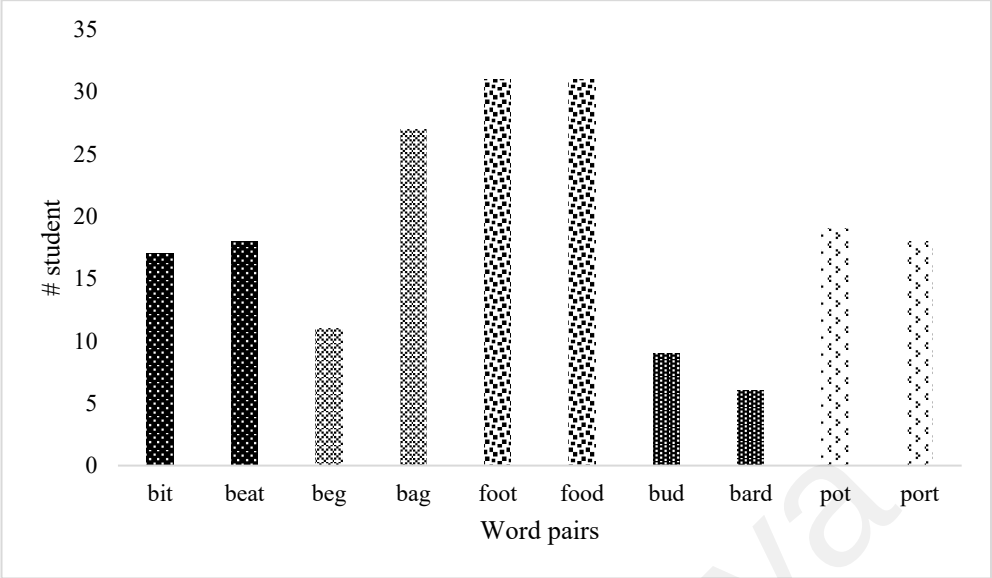

Figure 3.3: Comparison of word familiarity index in word pairs

CHAPTER 4: PRODUCTION OF THE ENGLISH MONOPHTHONGS BY ACEHNESE-INDONESIAN SPEAKERS

4.1 Introduction

This chapter presents the findings from the production of English monophthongs by the Acehnese-Indonesian speakers. The quality of the English monophthongs is first presented and followed by the comparison of typical English vowel contrasts in terms of their quality and length.

4.2 Quality of English Monophthongs

A total of 870 tokens (29 speakers x 10 words x 3 repetitions) generated from the English production task are presented in this section. The average formant frequencies and standard deviations for the F1 and F2 of each English monophthong produced by the Acehnese-Indonesian speakers can be observed in Table 4.1. The average formant frequencies ($\bar{x}$) are presented in both Hertz (Hz) and Bark scales accompanied by the standard deviation (SD). ED was also presented in the far-right column. The average distance of all Acehnese Indonesian English (Ach-IndE) vowels from the centroid [ɜ˞] is 2.62 Bark, and the average distances for males and females are 2.70 Bark and 2.54 Bark, respectively. This suggests that Ach-IndE vowel space is more peripheral or spread out than Singapore English (2.41 Bark) and Brunei English (1.82 Bark) but slightly smaller than Malaysian English (2.72 Bark) (Pillai, 2014). The larger vowel space of the Indonesian vowels (2.72 Bark) may cause the Ach-IndE vowels to be more peripheral (see section 5.3 ED of Indonesian vowels).

Table 4.1: Average first and second formant measurements for English monophthongs and ED

Vowel	Target Words	Gender	F1 (Hz)		F2 (Hz)		F1 (Bark)		F2 (Bark)		ED
			$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	
[i:]	beat	F	411	74	2626	338	3.9	0.7	14.7	0.8	3.76
		M	333	35	2315	139	3.2	0.3	14.0	0.4	3.62

[ɪ]	bit	F	390	56	2747	204	3.7	0.5	15.1	0.4	4.12
		M	337	44	2289	164	3.3	0.4	13.9	0.5	3.54
[e]	bet	F	625	69	2342	180	5.8	0.6	14.1	0.5	2.58
		M	524	59	2006	83	4.9	0.5	13.1	0.3	1.84
[æ]	bat	F	706	73	2170	248	6.4	0.6	13.6	0.7	2.14
		M	628	108	1786	233	5.8	0.8	12.3	0.9	0.79
[ʌ]	bud	F	795	92	1819	175	7.1	0.7	12.5	0.6	1.56
		M	559	169	1245	270	5.2	1.4	9.8	1.5	1.82
[ɑː]	pot	F	663	91	1277	116	6.1	0.7	10.1	0.6	1.43
		M	604	54	1116	131	5.6	0.4	9.2	0.7	2.33
[ɔː]	port	F	567	128	1146	169	5.3	1.0	9.3	1.0	2.24
		M	502	84	1002	89	4.7	0.7	8.5	0.6	3.20
[ʊ]	foot	F	465	56	1180	108	4.4	0.5	9.6	0.6	2.40
		M	404	47	1015	95	3.9	0.4	8.6	0.6	3.52
[uː]	food	F	481	45	1118	154	4.6	0.4	9.2	0.9	2.65
		M	397	57	998	102	3.8	0.5	8.5	0.6	3.65
[ɝ]*	bird	F	554	62	1596	220	5.2	0.5	11.6	0.9	0.65
		M	490	31	1382	164	4.6	0.1	10.6	0.8	1.50
Average		F									2.54
		M									2.70

n.b. *central

The distribution of English monophthongs produced by the speakers can be seen in Figure 4.1. The placement of the vowels in the vowel space was generally similar to American English vowels. The most salient features were that [iː] - [ɪ] and [ʊ] - [uː] were much closer together for the Acehnese-Indonesian English (Ach-IndE) in both males and females. The conflation of the vowel pair [iː] and [ɪ] was also reported in Malaysian English (Pillai, 2014; Pillai et al., 2010), Singaporean English (Deterding, 2003) and Brunei English (Sharbawi, 2006). The conflation of [ʊ] - [uː] was also observed in Brunei English but not in Malaysian English and Singapore English. The findings from the current study are also in stark contrast to the previous findings on Acehnese high school students studied by Fata, Fitrian, Mohammad, and Yusuf (2017). In their study, in addition to [iː] – [ɪ] and [ʊ] – [uː], the students also produced [ɛ] – [æ] and [ɒ] – [ɔ] vowel pairs close together.

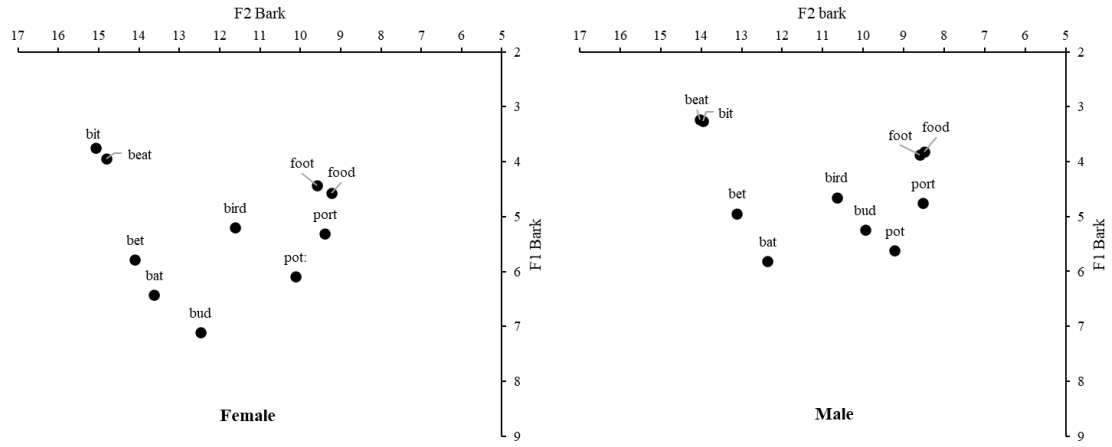

Figure 4.1: Vowel plot of English monophthongs

However, this visual assumption of differences among the vowel contrast would be further analysed using statistical analysis in the quality contrast section (4.3). Another interesting finding in the current study is that the [ʌ] in males was produced higher up and further back compared to females. A further explanation of this vowel for each speaker is presented in the mid vowel section (5.4.5).

4.2.1 Front Vowels

The scatter plot for Ach-IndE front vowel ([iː] in *beat*, [ɪ] in *bit*, [ɛ] in *bet* and [æ] in *bat*) is presented in Figure 4.2. It can be seen that both males and females do not separate the front vowel [iː] – [ɪ]. The conflation is more salient in males than females in that the two vowels were more crammed in males. The overlap is not as distinct as those reported in Malaysian English (Pillai, 2014, p. 70) but is mostly similar to Singapore English (Deterding, 2003, p. 9) and Brunei English (Sharbawi, 2006, p. 257). The standard deviation of [ɪ] is also higher in females compared to males (See Table 1.1), resulting in this vowel being more stretched in females. Some female speakers produced [ɪ] with a much higher F1 lowering the vowel further down and a higher F2 pushing the vowel higher up in the space.

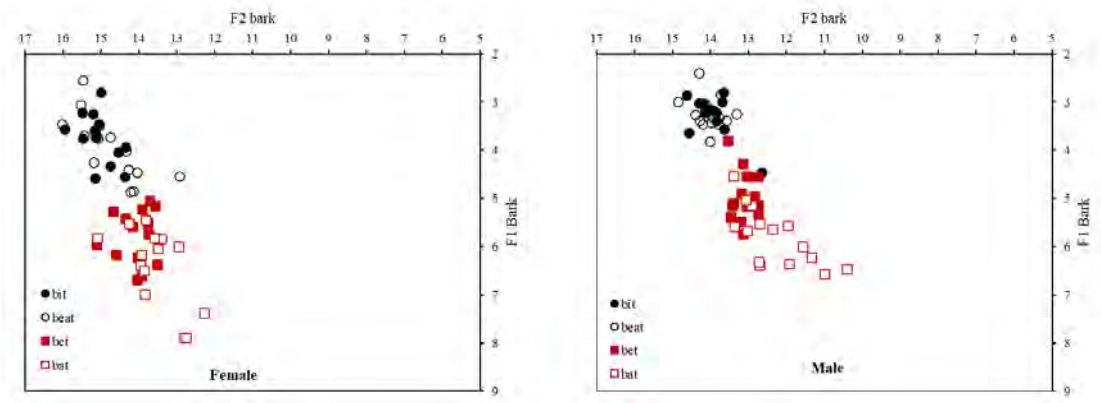

Figure 4.2: Front vowels

Categorical separation is also observed in [ɛ] – [æ] in both males and females, which is quite different from Malaysian English (Pillai, 2014; Pillai et al., 2010), Singapore English (Deterding, 2003) and Brunei English (Sharbawi, 2006). A previous study on Acehnese high school students reported conflation between [ɛ] – [æ] (Fata et al., 2017). However, based on the formant values and the placement of these two vowels in the vowel chart, the [æ] is placed lower than [ɛ]. Nine male and three female speakers also produced the [ɛ] vowel with higher F1 pushing the vowel further down the space. The higher F1 in some [ɛ] vowels indicates that this vowel might be produced similar to the Acehnese and Indonesian [a]. This is to be confirmed later when this vowel is plotted against the Acehnese and Indonesian [a] in section 2.4.4. A lower [æ] was also reported in Malaysian English (Pillai, 2014, p. 71; Pillai et al., 2010, p. 166) but not in Singapore English and Brunei English. Brunei English speakers (Sharbawi, 2006, p. 260) produced the [æ] vowel further back, while Singapore English speakers did not maintain any separation for the [ɛ] – [æ] vowels (Deterding, 2003, p. 10).

The box plot in Figure 4.3 illustrates the difference between F1 and F2 values of *bit* in male and females. It is visible in the box plot that the F1 for male speakers occupied the lower plot and was less varied compared to the female speakers except for the F2 values on the right. The box plot for F1 indicates that the MS8 speaker was the outlier and had

an average F1 value different from other speakers in the group. The MS8 had an average F1 value of 4.46 Bark, which was higher than the average. The average values of other speakers in the male group were below 4 Bark. The box plot for F2, on the other hand, had one female (FS6) and two males as outliers (MS7 & MS9). The FS6 had an average F1 value of 15.95 Bark, which was higher than the average. The average F2 values of other speakers in the female group were below 15.5 Bark. The MS10 had the lowest average F2 value at 2.85 Bark, while MS8 had the highest F2 value at 12.6 Bark.

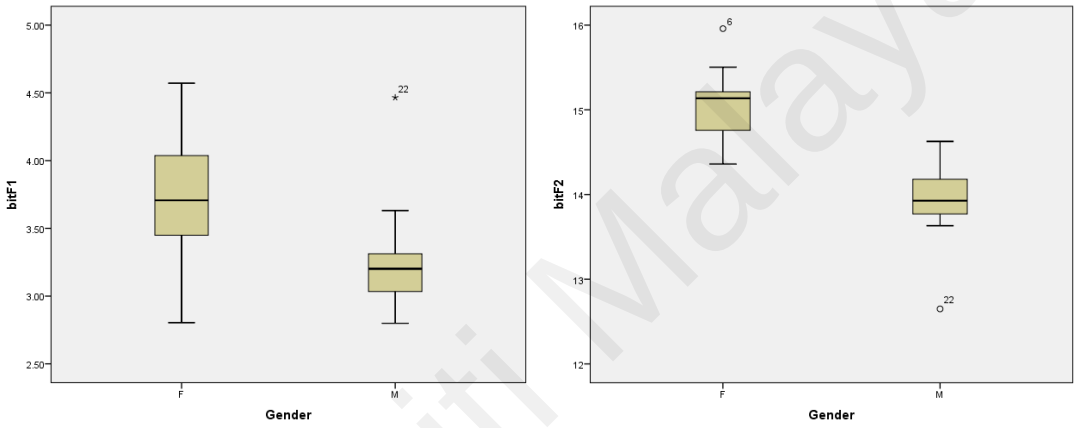

Figure 4.3: Box plot of F1 and F2 by gender for English vowel [ɪ] in *bit*

The box plot in Figure 4.4 illustrates the difference between F1 and F2 values of *beat* in male and females. It is visible in the box plot that the F1 for male speakers occupied the lower plot and is less varied compared to the female speakers. The mean for both groups was different, and the error bar did not show an overlap. It means that none of the data points in both groups were the same. The distribution of [iː] by gender is shown in Figure x. Obviously, there was no overlap between male and female speakers in the production of [iː], suggesting that they were produced differently.

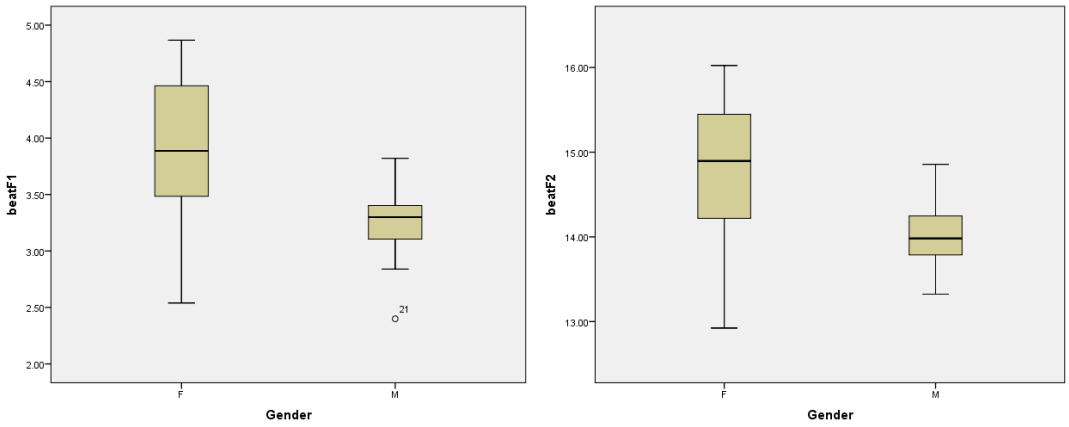

Figure 4.4: Box plot of F1 and F2 by gender for English vowel [iː] in *beat*

The box plot in Figure 4.5 illustrates the difference between F1 and F2 values of *bet* in male and females. It is visible in the box plot that the F1 for male speakers occupied the lower part and was equally varied with the female speakers. In contrast, for the F2 values on the right, the values were more spread in females than males. The average values of F2 in the male were mostly below 14 Bark while the females were above 13 Bark. No outliers were spotted for either F1 or F2 for the [e].

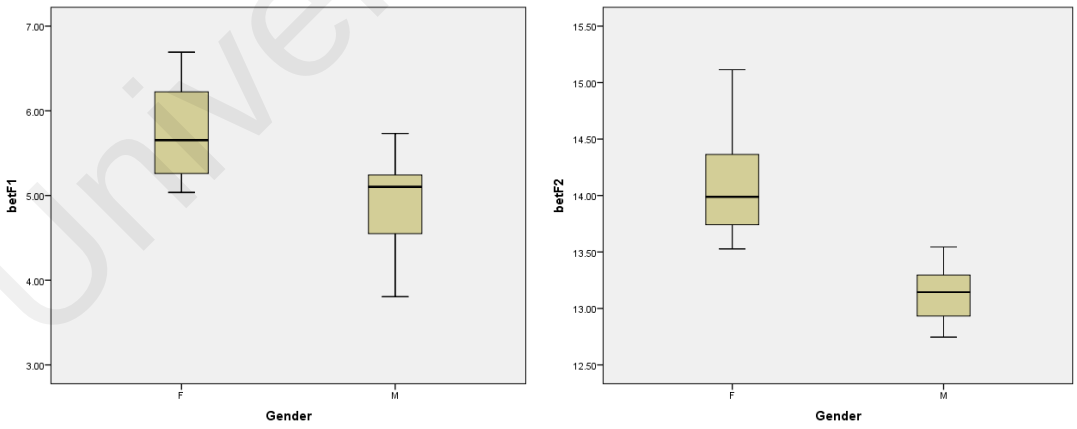

Figure 4.5: Box plot of F1 and F2 by gender for English vowel [ɛ] in *bet*

The box plot in Figure 20 illustrates the difference between F1 and F2 values of *bat* in male and females. It is visible in the box plot that the F1 and F2 for male speakers

occupied the lower half and were less varied compared to the female speakers. No outliers were spotted for either F1 or F2 for the [e]. The mean of each group indicated that there were differences, while the error bar did not overlap. It means that none of the data points in both groups are the same. The distribution of [æ] by gender is shown in Figure 4.6. There were mostly no overlaps between male and female speakers in the production of [æ], suggesting that they were produced differently.

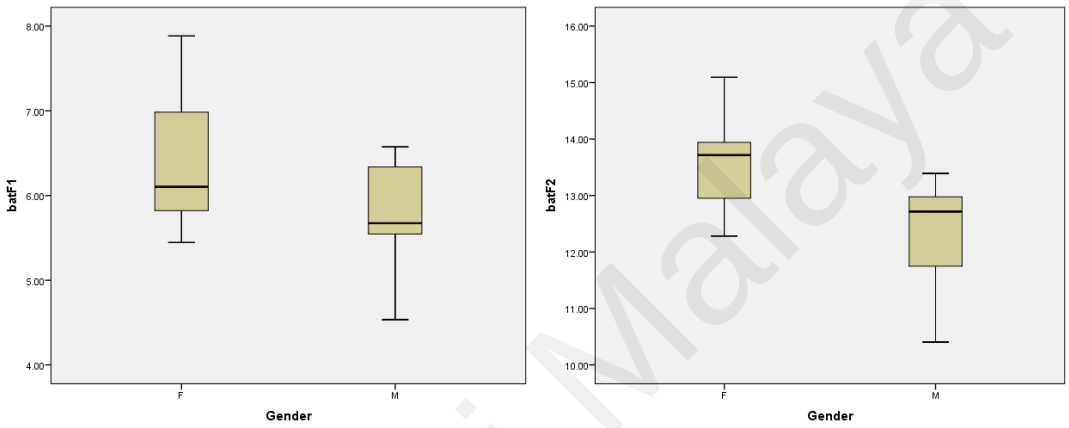

Figure 4.6: Box plot of F1 and F2 by gender for English vowel [æ]in bat

4.2.2 Mid Vowel

Figure 4.7 shows the scatter plot for the mid vowels [ʌ] in *bud* and [ɝ] in *bird* for the Ach-IndE speakers. Both males and females mostly produced this vowel pair apart in which [ʌ] was produced lower than [ɝ]. The most salient feature of the mid vowel was the sound of the post-vocalic r in *bird*, which suggests that Ach-IndE was rhotacised. As an effect of rhoticity, Hayward (2000) argues that the F2 of the vowel might be lowered. This impression analysis was further confirmed in the plot where the [ɝ] was produced with a much lower F2 making this vowel less fronted. R-coloured of the NURSE vowel was also reported in Brunei English (Sharbawi, 2006). Rhoticity was also found in two speakers of Malaysian English in Pillai (2014), in which the post-vocalic *r* was evident in *bird*, *board* and *bard*.

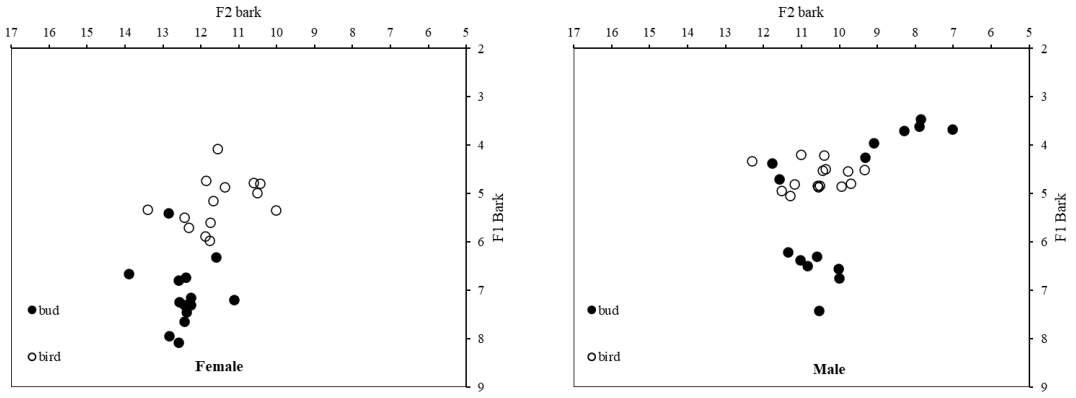

Figure 4.7: Mid vowels

The box plot in Figure 4.8 illustrates the difference between F1 and F2 values of *bud* in male and females. It is visible in the box plot that the F1 and F2 for male speakers occupied the lower half and were less varied compared to the female speakers. No outliers were evident for either F1 or F2 for the [e]. The mean of each group indicated there were differences, while the error bar did not overlap. This means that none of the data points in both groups were the same.

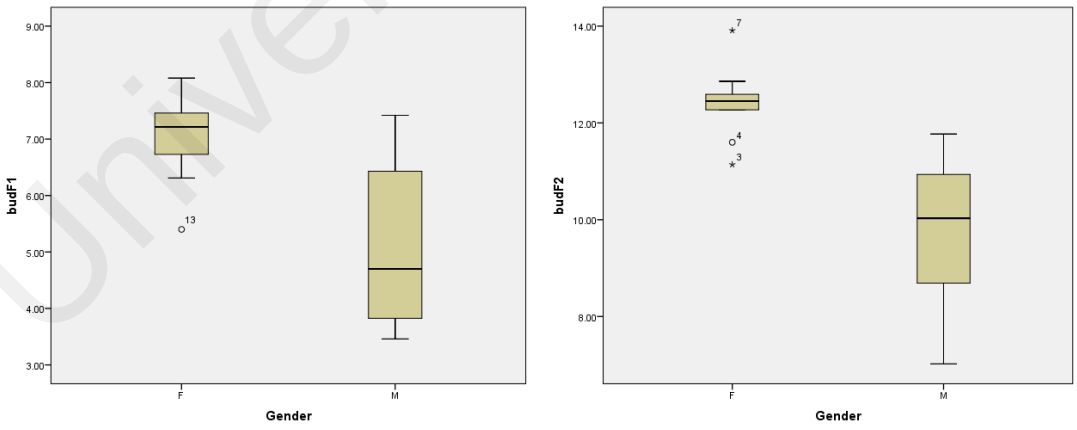

Figure 4.8: Box plot of F1 and F2 by gender for English vowel [ʌ] in *bud*

Another finding was that half of the male speakers produced the [ʌ] with a lower F1 and F2, indicating that it may be produced as an unrounded back vowel. Producing *bud* with

a lower F1 and F2 may suggest that they were produced as a high back vowel instead of as a mid-central vowel. It appears that the speakers pronounce the vowel in *bud* as [uː]. The standard deviation for the female speakers was also high at 0.9 Bark (See Table 1.1) compared to only 0.6 Bark for male speakers. This suggests that some speakers might produce this vowel as different vowels. This assumption was confirmed when this vowel was plotted against the English [ʊ] and [uː] in Figure 1.4 (left). Six male speakers produced the vowel in *bud* within the [ʊ] and [uː] space area, suggesting that these speakers relied on the written form when producing this vowel. Further confirmation can be seen when the English [ʌ] produced by males was compared to the Acehnese and Indonesian [u] in 2.4.5. Females, on the other hand, did not exhibit any extreme deviation in their production from the standard quality in their group. The production of the [ʌ] vowel in female speakers was mostly similar to those reported in the previous study (Fata et al., 2017).

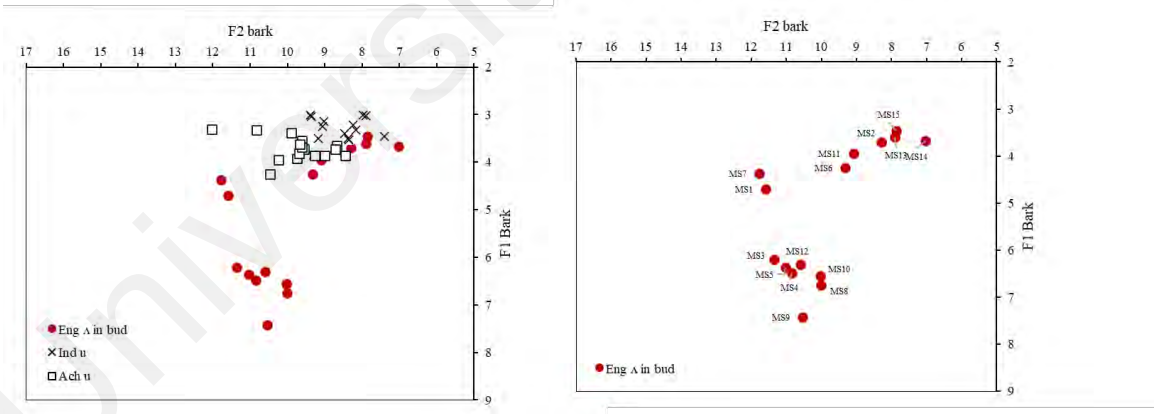

Figure 4.9: Comparison of English [ʌ] to English [ʊ] and [uː] (left) and English [ʌ] with speaker labels (right)

To confirm the production of *bud* as [ʌ] in male speakers, each speaker in the vowel plot was first labelled to determine which speakers produced the [ʌ] vowel as the back vowel. Figure 4.9 (right) labels all data points in the plot to determine the specific males whose

production was suspected as different from the rest of the males. The types of the sound produced by the speakers were then divided into three categories, that is, Category 1, Category 2, and Category 3, based on visual observation of the plot. The original tokens of each speaker were then revisited and classified to determine whether the word "bud" was produced as [ʌ] or [u] by each group of the speaker (Table 4.2). Thus, for this purpose, the analysis would be done based on the researcher listening carefully to the vowel produced in the targe words. Another coder was consulted to ensure consistency in the judgement of the vowel produced.

Table 4.2: Classification of English vowel [ʌ] variations produced by Ach-IndE speakers

Category 1 Assumed as [u]	MS2, MS6, MS11, MS13, MS14, and MS15
Category 2 Assumed as other vowels	MS1, MS7
Category 3 Assumed as [ʌ]	MS3, MS4, MS5, MS8, MS9, MS10, and MS12

In agreement with the initial assumption, the speakers in Category 1 produced the vowel in the word "bud" as [u] instead of [ʌ], suggesting that their production was influenced by the word's orthography "bud". The Indonesian pronunciation system expects the pronunciation of words to be based on their orthographies. One of the speakers, male speaker 1 (MS2), produced one of the three tokens as [bʌd] indicating that he was unsure of how to pronounce it. As for Category 2, mixed findings were discovered. MS1 produced the *bud* as [bɛd] in all three tokens, while MS7 initially produced the first token as [bəd] but he produced the rest of the tokens as [bɛd].

4.2.3 Back Vowels

Figure 4.10 shows the scatter plot for the back vowels [ɑː] in *pot*, [ɔː] in *port*, [ʊ] in *foot*, and [uː] in *food* for the Ach-IndE speakers. The first noticeable feature of the back vowel

was the conflation of [ʊ] – [uː] in both male and female speakers. A similar finding was reported in a previous study on Acehnese students (Fata et al., 2017, p. 594). Some degree of overlapping of this pair was also found in Brunei English (Sharbawi, 2006, p. 257), although less overlapping occurred in Malaysian English (Pillai, 2014, p. 72). Singapore English (Deterding, 2003, p. 8), on the other hand, was able to maintain the difference between the [ʊ] – [uː] vowels.

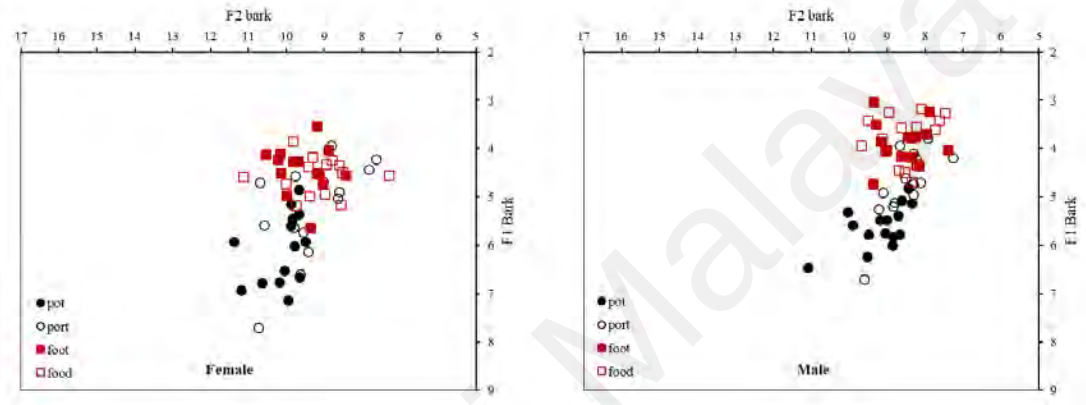

Figure 4.10: Back vowels

The second noticeable feature of the back vowel was the sounding of the post-vocalic *r* in *port* which suggests that in Ach-IndE this vowel was rhotacized. The F2 value of the [ɔː] vowel in Ach-IndE was lower, making this vowel less fronted compared to Malaysian English, Singapore English and Brunei English. This clearly indicates that the Ach-IndE vowels are closer to American English than to British English. R-colouring of the FORCE vowel was not reported in Malaysian English (Pillai, 2014; Pillai et al., 2010), Singapore English (Deterding, 2003) and Brunei English (Sharbawi, 2006). Having their roots in British English in, such features are to be expected in the English spoken Malaysia, Singapore and Brunei.

The box plot in Figure 4.11 illustrates the difference between F1 and F2 values for *pot* in males and females. It is visible in the box plot that the F1 for male speakers occupied the

lower half and was less varied compared to the female speakers. As for the F2, female speakers tended to produce it at a higher rate compared to males, while variation was small. Two females (FS7 and FS11) and one male (MS5) produced the F2 values differently from their respective groups. The F2 value of FS7 (11.37 Bark) and FS11 (11.19 Bark). The F2 value by MS5 (11.09 Bark). This suggests that this speaker produced the vowel with a quality different from other male speakers. However, the mean of each group indicated differences while the error bar did not overlap. It means that none of the data points in both groups were the same.

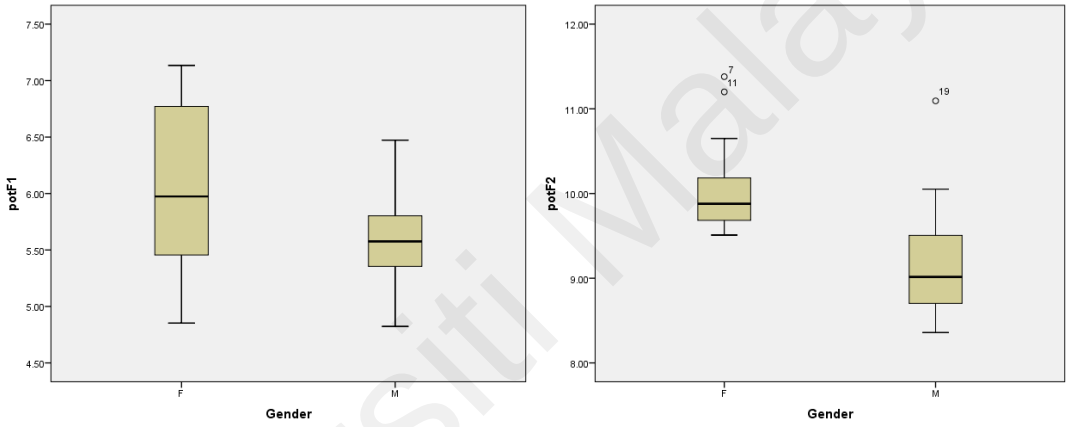

Figure 4.11: Box plot of F1 and F2 by gender for English vowel [ɑː] in *pot*

The box plot in Figure 4.12 illustrates the difference between the F1 and F2 values for *port* in male and females. It is visible in the box plot that the F1 for male and female speakers were almost comparable, while some degree of overlap at the lower end is apparent. As for the F2, female speakers tend to produce it at a higher rate compared to males while variation was higher in females than in males. One female (FS6) and one male (MS9) produced the F1 values differently from their respective group, while no outlier was observed in F2. The F1 value of FS6 (10.7 Bark) and MS9 (9.59 Bark) was. However, the mean of each group indicated differences, while the error bar showed minimal overlap. Few data points in both groups were the same in F2 but not in F1.

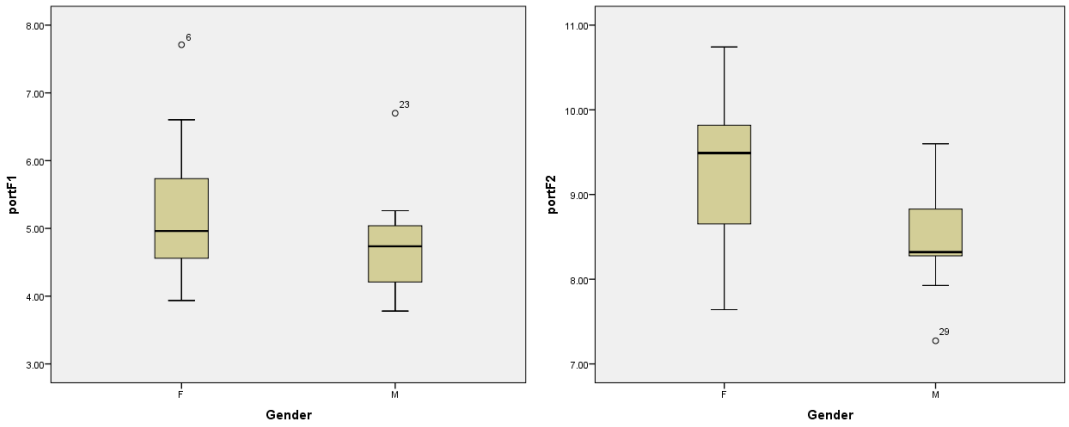

Figure 4.12: Box plot of F1 and F2 by gender for English vowel [ɔː] in port

The box plot in Figure 4.13 illustrates the difference between the F1 and F2 values for *foot* in male and females. It is visible in the box plot that the F1 for male and female speakers were almost comparable, while some degree of overlap at the lower end was apparent. As for the F2, female speakers tended to produce it at a higher rate than males, while variation was higher in females than in males. One female (FS11) produced the F1 values differently from other females in the group. The F1 value of FS11 was 9.3 Bark. No outlier was spotted in the F2 plot, both for males and females. The mean of each group indicated differences, while the error bar showed minimal overlap. Few data points in both groups were the same in F2 but not in F1.

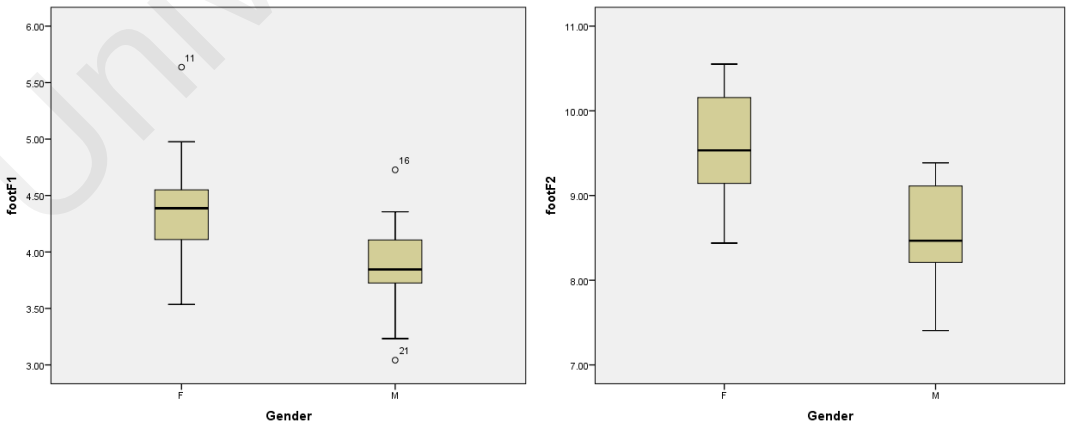

Figure 4.13: Box plot of F1 and F2 by gender for English vowel [ʊ] in *foot*

The box plot in Figure 4.14 illustrates the difference between the F1 and F2 values for *food* in male and females. It is visible in the box plot that the F1 for male and female speakers were almost comparable while some degree of overlap at the lower end was apparent. As for the F2, female speakers tended to produce it at a higher rate than males, while variation was comparable in both females and males. No outlier was spotted in either F1 and F2 plots in both groups. The mean of each group indicated differences, while the error bar showed minimal overlap.

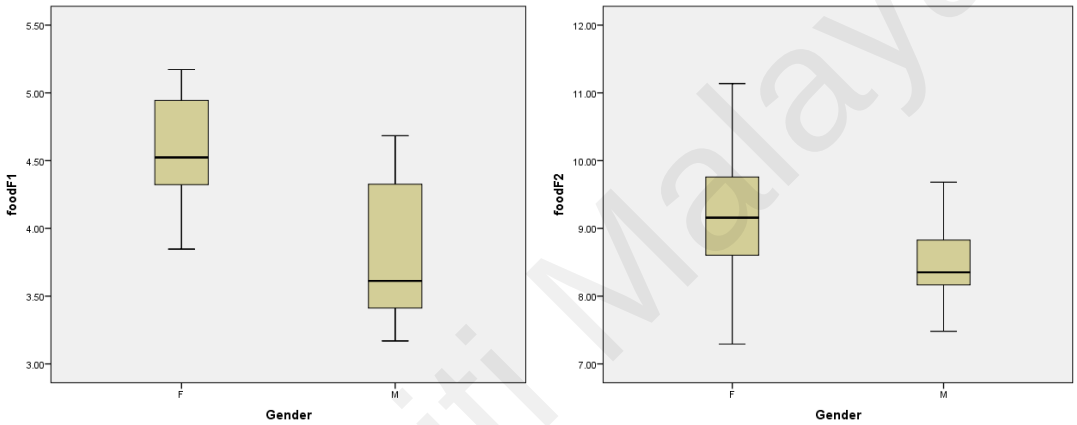

Figure 4.14: Box plot of F1 and F2 by gender for English vowel [u:] in food

4.3 Vowel Contrast

This section focuses on the quality contrast of five English vowel pairs. The five English pairs can be observed in

Table 4.3: Classification of English vowel variation

Vowel contrasts	Target words
[ɪ] - [iː]	bit - beat
[ɛ] - [æ]	beg - bag
[ʊ]- [uː]	foot - food
[ʌ] - [ɑː]	bud - pot
[ɑː] - [ɔː]	pot - port

4.3.1 Vowel Contrast for Bit-Beat

The distribution of *bit - beat* can be observed in Figure 4.15. Looking at the scatter plot, it is apparent that the production of both vowels overlapped, indicating that they were produced similarly. To confirm this assumption, a *t*-test was conducted. A paired sample *t*-test showed no significant difference in the average F1 ($t(28) = 0.659$, $p = 0.515$) and F2 ($t(28) = -1.143$, $p = 0.263$) between [ɪ] and [iː] vowel contrast. This result suggested that the Ach-IndE speakers did not distinguish the quality of English *bit - beat* in their production. This finding confirmed the findings in the previous study that Acehnese-Indonesian speakers were not able to distinguish the quality of this vowel pair in their production (Fata et al., 2017). Lack of categorical separation for this pair was also salient in Javanese and Sundanese Indonesian learners of English (Perwitasari et al., 2016). The conflation of the English vowels in beat and bit, had also been reported as an emerging feature in countries neighbouring Indonesian, particularly Malaysia (Pillai, 2014; Pillai et al., 2010), Singapore (Deterding, 2003, 2007) and Brunei (Sharbawi, 2006).

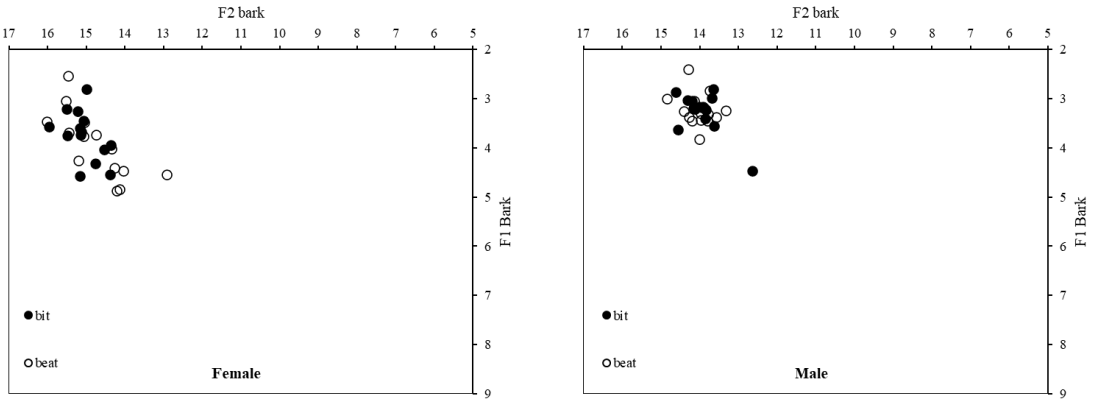

Figure 4.15: Comparison of English vowel pair in *bit-beat*

4.3.2 Vowel Contrast for Bet and Bat

The scatter plot of *bet - bat* can be observed in Figure 4.16. The scatter plot shows that the female speakers seemed to conflate the two vowels while the male speakers did not. Most male speakers produced the [æ] vowel lower and further back than the [ɛ] vowel, indicating that the two vowels were distinct. To confirm whether the pair were conflated, a *t*-test was conducted. A paired sample *t*-test showed significant difference in the average F1 ($t(28) = -7.101$, $p < 0.001$) and F2 ($t(28) = 4.908$, $p < 0.001$) between *bet - bat* vowel contrast. It means that there was a significant difference in the production of *bet - bat* vowel contrast both in terms of F1 and F2 by the Acehnese-Indonesian speakers.

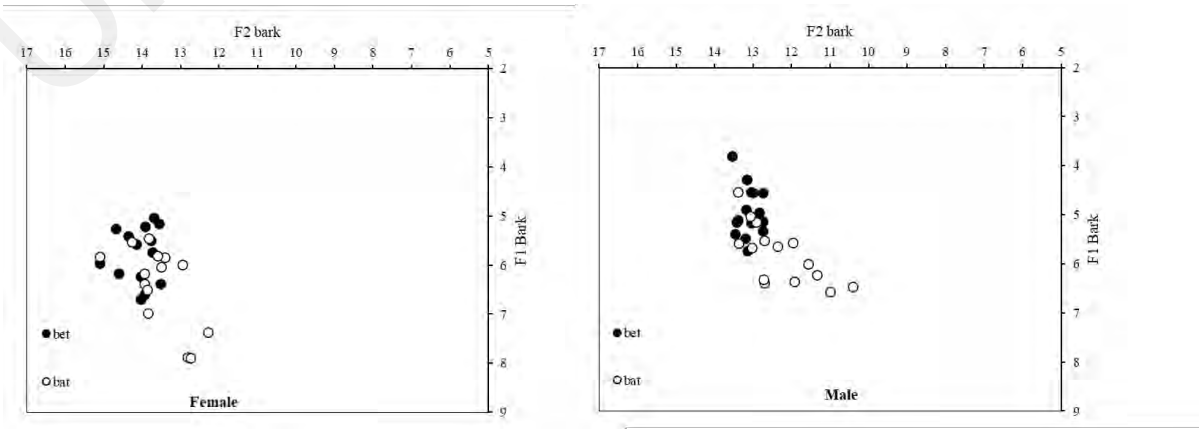

Figure 4.16: Comparison of English vowel pair in *bet-bat*

It is interesting to find that the *bet - bat* pair had a distinctive quality in the current study, as a previous study on Acehnese-Indonesian speakers did not yield the same findings (Fata et al., 2017). Javanese and Sundanese Indonesian speakers also did not distinguish this vowel pair in a previous study (Perwitasari et al., 2016). The lack of categorical separation in the *bet - bat* pair was also a salient feature in Malaysian English (Pillai, 2014; Pillai et al., 2010), Singapore English (Deterding, 2003) and Brunei English (Sharbawi, 2006). The distinctive quality of vowels in the *bet-bat* pair does not necessarily mean that the current speakers produced *bat* with the [æ] vowel, as it was later confirmed in 2.4.4 when this vowel was compared to their production of the Acehnese and Indonesian [a].

4.3.3 Vowel Contrast for Bud-Pot

The scatter plot of *bud-pot* can be observed in 4.17. The scatter plot make it apparent that the production of both vowels did not overlap, indicating that they were produced differently. The spread of the vowel in the word *bud* was more apparent in males than females. Female speakers consistently produce *bud* at the low mid location, while the male speakers are inconsistent in their production of this vowel. The vowel in the word *bud* for the male speakers was spread from the back to front mid position of the vowel space and all the way to the bottom mid location of the vowel space, which suggests that they may have produced the vowel in *bud* with a range of vowel quality. A paired sample *t*-test showed no significant difference in the average F1 ($t(28) = 1.310$, $p = 0.201$) but F2 ($t(28) = 6.007$, $p < 0.001$) between vowels in the two words suggesting that they were produced similarly. It means that the significant difference in the production of *bud-pot* vowel contrast was only observed in terms of F2 but not in F1. However, it is important to note that this categorical separation might be attributed to the male speakers producing the vowel in *bud* similar to the English [ʊ] and [uː] as shown in the previous section on the mid central vowel (4.2.2). However, it does not necessarily imply that Acehnese-

Indonesian speakers produced the vowel in *bud* as the [ʌ] vowel despite [ʌ] existing in the Acehnese vowel system. A comparison of the vowel in *bud* to equivalent Acehnese and Indonesian vowels can be seen in 2.4.5.

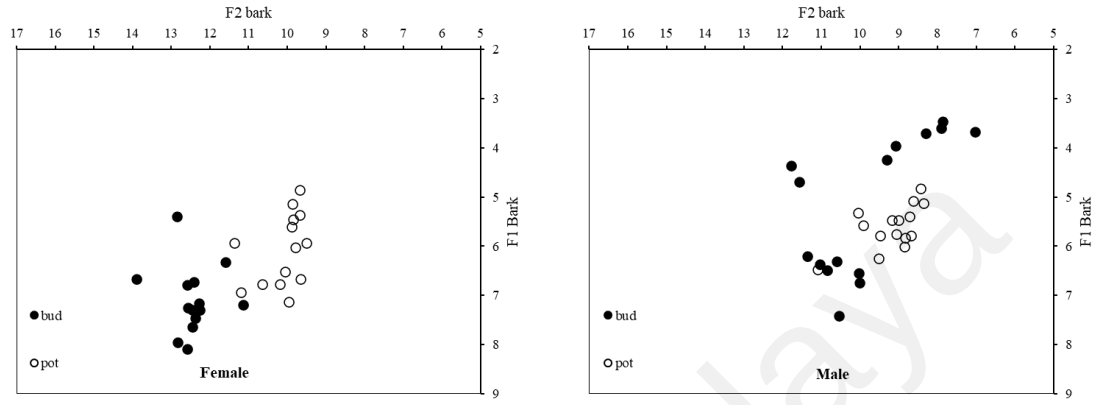

Figure 4.17: Comparison of the English vowel pair in *bud-pot*

The categorical separation of this pair was also reported in a previous study by Acehnese-Indonesian speakers (Fata et al., 2017). Javanese and Sundanese Indonesian speakers have also been reported to separate this vowel pair to some degree (Perwitasari et al., 2016). This finding also suggests that Acehnese-Indonesian speakers shared the same feature as Javanese and Sundanese speakers for this vowel pair. This separation was not observed in Malaysian English (Pillai, 2014, p. 72), and Singapore English (Deterding, 2003, p. 8), but in Brunei English (Sharbawi, 2006, p. 253). However, the Brunei English [ɑː] was higher than [ʌ], while in the current study [ʌ] is higher than [ɑː]. For this vowel pair, Indonesian English did not share the same quality feature with Malaysian English and Singapore English but to some extent, the overall position was similar to Brunei English.

4.3.4 Vowel Contrast for Pot-Port

The scatter plot of the vowels in the words *pot* and *port* can be observed in Figure 4.18. The scatter plot makes it apparent that the production of both vowels did not overlap, indicating that they were produced differently. To confirm this assumption, a *t*-test was conducted. A paired sample *t*-test showed that there was a significant difference in the average F1 ($t(28) = 5.696$, $p < 0.001$) and F2 ($t(28) = 6.059$, $p < 0.001$) between *pot* and *port* vowel contrast. This means that there was a significant difference in the production of *pot* - *port* vowel contrast both in terms of F1 and F2 by the Acehnese-Indonesian speakers.

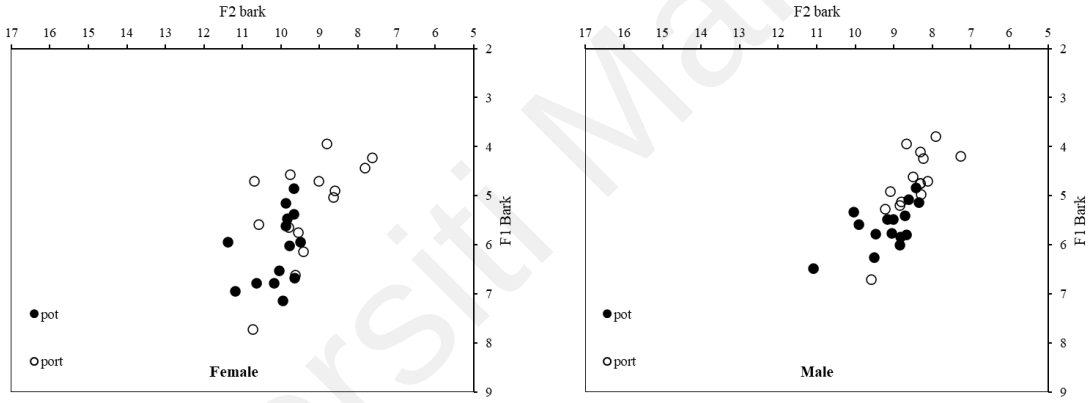

Figure 4.18: Comparison of English vowels pair in *pot-port*

The separation of this pair was not found in the previous study by Indonesian speakers. Acehnese-Indonesian speakers in the study by Fata et al. (2017) did not distinguish this vowel pair and neither did the Javanese and Sundanese Indonesian speakers in the study by Perwitasari et al. (2016). One possible explanation for this is that the previous studies used words that did not have the post-vocalic r. This further suggests that the Acehnese-Indonesian speakers produced r-coloured vowels in *port* which is apparent from the vowel being pushed back and resulting in lower F2 values. This [ɔː] vowel was also previously shown to be positioned near the back vowels [ʊ] and [uː] (see Figure 4.10). Complete neutralisation between the quality of this vowel was also found in Malaysian English

(Pillai, 2014; Pillai et al., 2010), Singapore English (Deterding, 2003) and Brunei English (Sharbawi, 2006, 2012).

4.3.5 Vowel Contrast for Foot-Food

The scatter plot of the vowels in *foot-food* vowels can be observed in Figure 4.19. By looking at the scatter plot, it is apparent that the production of both vowels overlapped, indicating that they were produced similarly. To confirm this assumption, a *t*-test was conducted. A paired sample *t*-test shows no significant difference in the average F1 (*t(28)* = -0.530, *p* = 0.601) and F2 (*t(28)* = -1.825, *p* = 0.079) between the *foot* and *food* vowel contrast. It means that there was no significant difference in the production of the *foot-food* vowel contrast in terms of F1 and F2 by the Acehnese-Indonesian speakers.

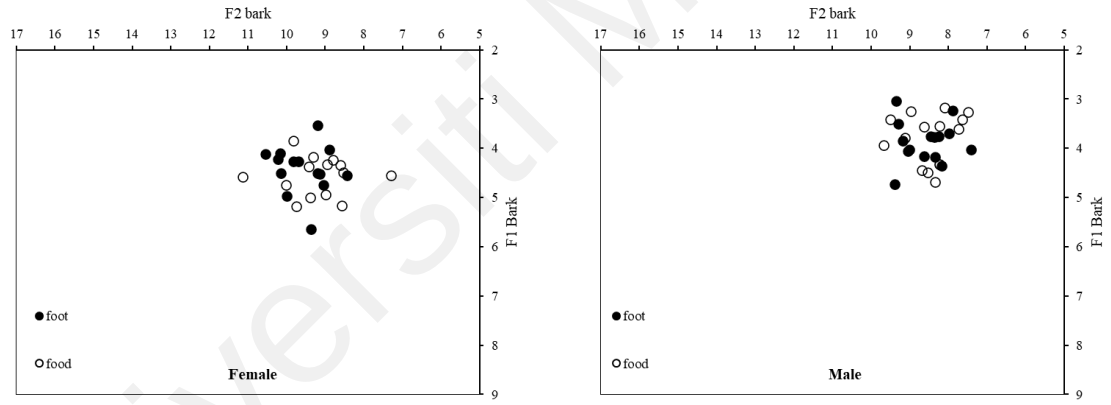

Figure 4.19: Comparison of English vowel pair in *foot-food*

The lack of categorical separation was mentioned in a previous study on Acehnese-Indonesian speakers (Fata et al., 2017). The same finding was also reported in Javanese and Sundanese Indonesian speakers (Perwitasari et al., 2016). This suggests that the Indonesian speakers could not distinguish this vowel pair in their production. In terms of neighbouring countries, the lack of categorical separation was a salient feature of Brunei English (Sharbawi, 2006, 2012). However, the conflation of this pair did not occur in Malaysia English (Pillai, 2014; Pillai et al., 2010) and Singapore English (Deterding, 2003) despite the three countries being former British colonies.

4.3.6 Durational Contrast

Table 4.4 provides the mean duration of the ten vowels produced by the male and female speakers. Overall, the vowel in *bird* had the longest duration at 158 milliseconds. The second vowels with the longest duration were *port* and *bat* at 144 and 143 milliseconds, respectively. The longer duration of *bird* and *port* may be influenced by the presence of the post-vocalic r, which was shown to be rhotic in the previous section. The vowels in *foot* and *bit* had the shortest duration at 108 milliseconds. Most vowels had a duration of between 114 and 141 milliseconds. The vowels in *food* and *beat* had a duration of 114 and 124 milliseconds, while vowels in *bey*, *bud*, and *pot* had a duration of between 132-141 milliseconds.

Table 4.4: Duration of English monophthongs in a millisecond

Target Word	Vowel	Female		Male	
		$\bar{x}$	SD	$\bar{x}$	SD
beat	[iː]	124	52	115	26
bit	[ɪ]	108	31	109	23
bet	[ɛ]	135	33	124	27
bat	[æ]	143	49	130	27
bud	[ʌ]	141	34	123	23
pot	[ɑː]	132	32	120	22
port	[ɔː]	144	39	140	25
foot	[ʊ]	108	31	101	21
food	[uː]	114	45	121	36
bird	[ɝ]	158	36	138	28

The comparison of vowel duration of both males and females are projected in Figure 4.20. It is clear from the figure that the vowels in *bird* and *port* are the longest vowel while [ɪ] and [ʊ] are the shortest. Both males and females consistently produced each vowel at a comparable duration except for *food*. Female vowels are also consistently higher than males, except for the vowel in *bit* and *food*. Female speakers produced the vowel in *bit* at a similar duration while the vowel in *food* with a slightly shorter duration than the males speakers. To confirm if the current speakers exhibited any durational contrast, a paired

pair *bit – beat, bud – pot, pot – port,* and *foot – food*. The result indicated that no significant difference in duration was recorded for *bud – pot* ($t(28) = 2.637$, $p = 0.013$) and *foot – food* ($t(28) = -2.120$, $p = 0.043$) yet a significant difference was observed for *bit – beat* ($t(28) = -2.314$, $p = 0.028$) and *pot – port* ($t(28) = -4.332$, $p = 0.001$). In other words, the Acehnese-Indonesian speakers maintain vowel duration for *bit – beat* and *pot – port*, but not for *bud - pot* and *foot - food*.

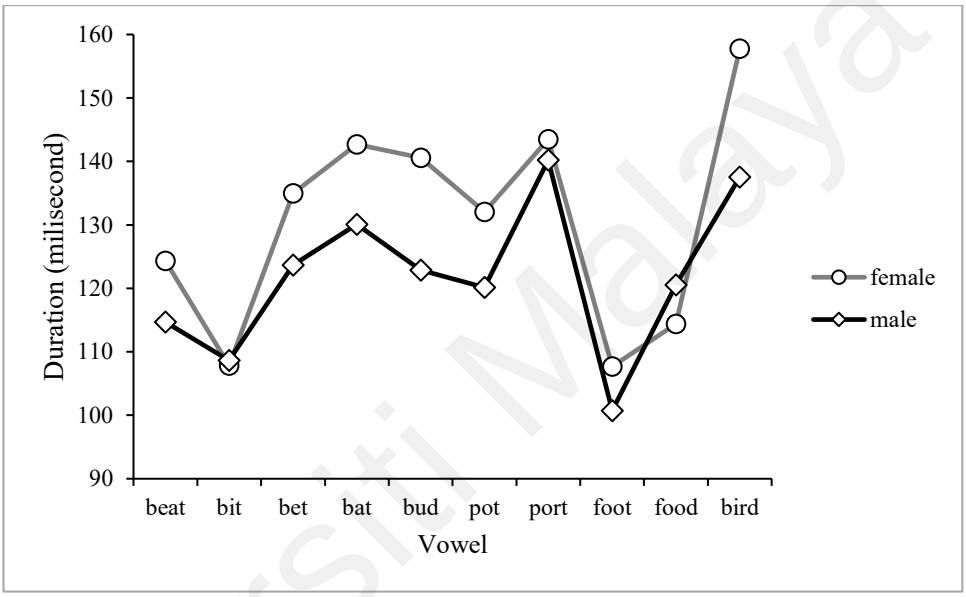

Figure 4.20: English vowel duration comparison based on gender

4.4 Conclusion

This chapter has presented the findings from the production data. The overall quality of Acehnese-Indonesian English monophthongs has been presented, followed by a detailed description of each vowel in terms of front, mid and back vowels. Five traditional vowel pairs were compared in terms of quality and durational contrast, and statistical analyses were made to determine their significant differences. The findings showed that Ach-IndE vowel space was more peripheral than in SgE (Deterding, 2003) and BrunE (Sharbawi, 2006). The overall quality of Ach-IndE was similar to the previous study on Acehnese

speakers (Fata et al., 2017) and Javanese and Indonesian speakers (Perwitasari et al., 2016). In terms of vowel contrast, Ach-IndE speakers maintain categorical separation for *bet – bat, bud – pot,* and *pot – port* but not for *beat – bit and foot - food*. In term of durational contrast, the Ach-IndE speakers maintain vowel duration for *beat – bit* and *pot – port* but not for *bud - pot* and *foot - food.*

CHAPTER 5: COMPARISON OF ENGLISH MONOPHTHONG AND EQUIVALENT ACEHNESE-INDONESIAN VOWELS

5.1 Introduction

This section compares the quality of English monophthongs with equivalent Acehnese-Indonesian monophthongs produced by Acehnese-Indonesian bilingual speakers of English. Prior to comparison, the quality of Acehnese and Indonesian monophthong vowels is presented based on the average values of F1 and F2 produced by all speakers. Later, the F1 and F2 values of English vowels were compared against the F1 and F2 values of equivalent Acehnese and Indonesian vowels. Repeated measure ANOVA was used to determine whether English vowels were significantly different from equivalent Acehnese and Indonesian vowels.

5.2 Quality of Acehnese Monophthongs

The Acehnese monophthong vowels, words used to elicit the vowel, average formant frequencies, and standard deviations for the F1 and F2 of each vowel produced by Acehnese-Indonesian speakers can be observed in Table 5.1. Average formant frequencies are presented in both Hertz (Hz) and Bark scales. EDs (in Bark), or ED, are also presented in the far-right columns. ED was calculated to show how spread out the Acehnese monophthongs compared to previously reported study of speakers from another part of Aceh. The average distance of all Acehnese vowels from the centroid [ə] is 2.26 Bark, and the average distances for males and females are 2.14 Bark and 2.37 Bark, respectively. This suggests that the West Acehnese vowel space is more peripheral or spread out than the North Acehnese vowel (2.12 Bark) reported in the previous study (Yusuf, 2013, p. 113).

Table 5.1: Average first and second formant measurements for Acehnese monophthongs

Vowel	Target Words	Gender	F1 (Hz)		F2 (Hz)		F1 (Bark)		F2 (Bark)		ED
			x̄	SD	x̄	SD	x̄	SD	x̄	SD	
[i]	bit	F	394	38	2711	270	3.8	0.2	15.0	1.0	4.04
		M	332	20	2253	275	3.2	0.2	13.8	1.0	3.48
[e]	pét	F	484	25	2388	75	4.6	0.2	14.1	0.2	2.88
		M	403	25	2199	78	3.9	0.2	13.7	0.2	2.95
[ɛ]	cèt	F	652	23	2247	119	6.0	0.2	13.8	0.4	2.31
		M	546	23	2000	122	5.1	0.2	13.1	0.4	1.73
/ɯ/	peut	F	453	37	1791	149	4.3	0.3	12.4	0.7	1.73
		M	394	38	1443	155	3.8	0.3	10.9	0.7	2.14
[ə]*	tet	F	526	30	1688	242	5.0	0.3	11.9	1.2	0.95
		M	437	31	1458	242	4.2	0.3	10.9	1.2	1.77
[ʌ]	göt	F	643	23	1929	354	5.9	0.2	12.8	1.4	1.27
		M	539	24	1749	355	5.1	0.2	12.1	1.4	0.96
[a]	pat	F	526	40	1688	84	5.0	0.3	11.9	0.4	0.95
		M	437	41	1458	87	4.2	0.3	10.9	0.4	1.77
[ɔ]	pôt	F	683	36	1284	60	6.2	0.3	10.1	0.3	1.43
		M	584	36	1156	57	5.4	0.3	9.4	0.3	2.10
[o]	cop	F	495	30	1064	105	4.7	0.3	8.9	0.6	2.85
		M	443	28	1002	106	4.2	0.3	8.5	0.6	3.41
[u]	cut	F	453	27	1361	173	4.3	0.3	10.5	0.9	1.83
		M	386	28	1212	179	3.7	0.3	9.7	0.9	2.78
Average		F									2.14
		M									2.37

n.b. *central

Figure 5.1 provides the scatter plot for Acehnese monophthong vowels produced by West Aceh speakers. The West Aceh dialect is one of the major dialects in Aceh for the people who live along the west coast of Aceh. The overall production of the West Acehnese dialect monophthongs is similar to the Pase dialect described in Pillai and Yusuf (2012). However, some minor differences from the West Aceh dialect were found. First, the length of the vowel space in the West Acehnese speakers was slightly smaller, while the height was slightly bigger. Second, the position of [ʌ] was higher while the [ɛ] was located lower in the space compared to the Pase dialect which was in reverse to what was found in the Pase dialect. The [ə] was produced also toward the back in the West Aceh

dialect compared to the Pase dialect. This position of [ə] being pushed back was also reported in the West Acehnese dialect by (Tanzir Masykar, Roni Agusmaniza, et al., 2021). Both males and females also produced the vowels roughly in a comparable position. As expected, vowels produced by the female speakers were lower in space due to having higher frequencies of F1 values. On the other hand, the males' vowels were higher in the vowel space due to having smaller F1 values. The distribution of each vowel was equally similar in both male and female speakers. Two obvious differences were found for [ɯ] – [ə] and [i] – [e]. The relative position of [ɯ] – [ə] and [i] – [e] were much closer together in males than in females.

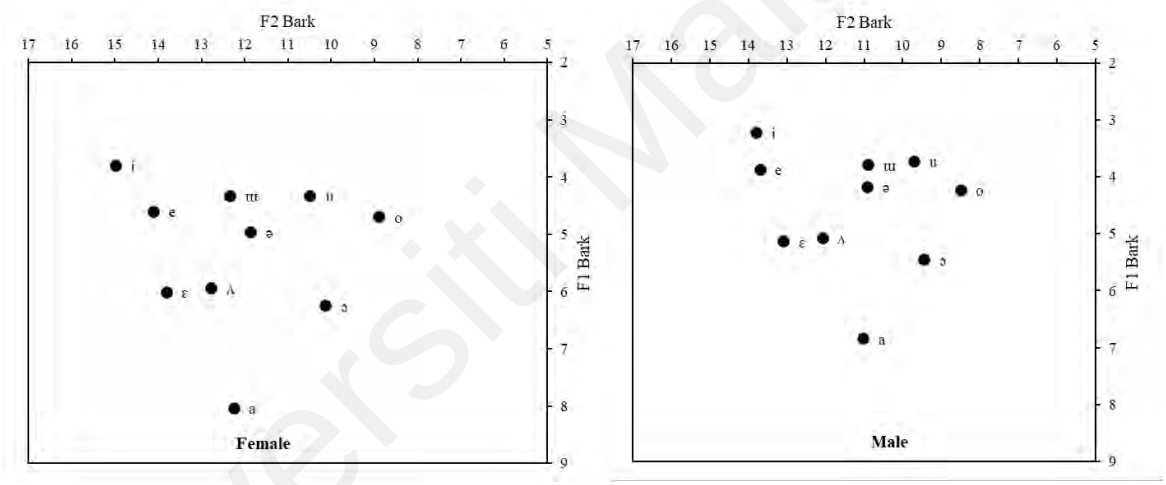

Figure 5.1: Vowel plot of Acehnese monophthongs

The production of some of the front vowels [i] and [e] were conflated to some extent in the middle, but the speakers mostly distinguished the vowels well. The front vowel [ɛ] was produced lower than the front vowel [i] and [e]. Separation was maintained for the vowel [ɛ] from the vowel [e]. As for the mid vowels [ə] and [ɯ], the speakers were also able to distinguish the two vowels, while some minor conflation could be observed. However, the vowel [ə] occupied a much larger area spanning from the front vowel [ɛ] all the way back to the vowel [o]. It was not clear why such a case happened, as this was

not recorded in the previous study of the same dialect (Tanzir Masykar, Roni Agusmaniza, et al., 2021). The production of the mid vowel [ʌ] also spread toward the front vowel [ɛ], back vowel [ɔ] and mid vowel [ə]. Such variation was also reported in another study by (Tanzir Masykar, Roni Agusmaniza, et al., 2021). Their study showed that the [ʌ] vowel was realized with variations depending on where the vowel sat in the Acehnese words. They argued that "the vowel [ʌ] was realised as [ə], [ɛ], and [ɔ] by the people in Aceh Barat." Such a variation of [ʌ] did not happen in the Pase dialect.

5.3 Quality of Indonesian Monophthong

The Indonesian monophthong vowel, words used to elicit the vowels, the average formant frequencies, and standard deviations for the F1 and F2 of each vowel produced by Acehnese-Indonesian speakers can be observed in Table 5.2. Average formant frequencies were presented in both Hertz (Hz) and Bark scales. The Hz scale is used to describe the properties of each monophthong vowel relative to the others, while the Bark scale is used to plot the vowel into the vowel space area. EDs (in Bark), or ED, are also presented in the far-right column. The average distance of all Indonesian vowels from the centroid [ə] was 2.72 Bark, and the average distances for males and females were 2.63 Bark and 2.82 Bark, respectively. This suggests that the Indonesian vowel space of the current speakers is more peripheral or spread out than their English and Acehnese vowels.

Table 5.2: Average first and second formant measurements for Indonesian monophthongs

Vowel	Target Words	Gender	F1 (Hz)		F2 (Hz)		F1 (Bark)		F2 (Bark)		ED
			x̄	SD	x̄	SD	x̄	SD	x̄	SD	
[i]	bibit	F	405	31	2590	88	3.9	0.3	14.7	0.3	3.73
		M	351	28	2214	92	3.4	0.3	13.7	0.3	3.32
[ɛ]*	bebek	F	567	25	2389	134	5.3	0.2	14.2	0.4	2.76
		M	495	26	2045	138	4.7	0.2	13.2	0.5	2.07
[ə]	tetap	F	818	62	1582	55	7.2	0.5	11.5	0.3	1.35
		M	702	64	1430	51	6.4	0.5	10.9	0.2	0.85
[a]	babat	F	792	70	1818	206	7.1	0.6	12.5	1.1	1.57
		M	622	71	1346	209	5.8	0.6	10.4	1.1	1.11
[ɔ]	bobok	F	646	82	2003	210	5.9	0.7	11.6	1.0	0.13
		M	600	83	1292	215	5.6	0.7	10.1	1.1	1.41
[o]	bobot	F	458	25	895	46	4.4	0.2	7.8	0.3	3.99
		M	403	24	822	38	3.9	0.2	7.3	0.3	4.65
[u]	bubut	F	384	27	786	67	3.7	0.2	7.0	0.5	4.99
		M	344	23	722	45	3.3	0.2	6.5	0.3	5.56
Average		F									2.86
		M									3.02

n.b. *central vowel

The distribution of Indonesian monophthong vowels produced by Acehnese-Indonesian speakers can be seen in Figure 5.2. The most noticeable feature of the Indonesian vowels produced by these speakers was that the position of the [ə] vowel was reversed with [a] compared to what was previously described by Zanten and Heuven (1984) and Soderberg and Olson (2008). In their description of the Indonesian vowels, the position of [ə] was higher than the [a] vowel, while in the current study, it was the [a] that was higher. The male and female speakers also produced these two vowels at slightly different locations. The [a] vowel was slightly higher and more to the front than [ə] in females, while in males the [a] vowel was much higher and more to the back than [ə]. Another salient feature was that the male speakers produced the [a] vowel closer to [ɔ] while the female speakers produced them a little bit further from each other. The front and back vowels were similar to what was described in the previous study (Soderberg & Olson, 2008; Zanten & Heuven,

1984). The vowel [ə] was lower and more to the back than [e], while the unrounded [u] was higher and more to the back than the rounded [o]. However, the distance between [i] and [e] was double than the distance between [u] and [o].

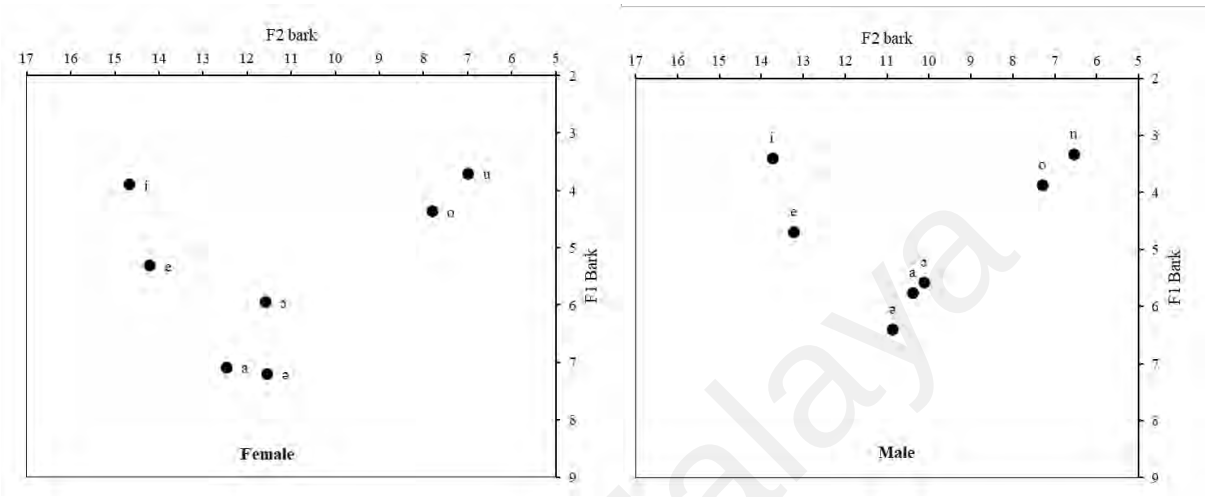

Figure 5.2: Vowel plot of Indonesian monophthongs

5.4 English Vowels Similar to Acehnese and or Indonesian

Some English vowels were (in terms of phonetic symbols) similar to Acehnese and Indonesian vowels. This similarity was considered based on identical IPA symbols used to represent English, Acehnese and Indonesian sounds. The use of phonetic symbols to compare the quality of sounds in a different language was also used in previous studies (Bohn & Flege, 1992; Flege, 1987), and more recently in the context of Indonesia (Perwitasari et al., 2016). Five English vowels had comparable sounds in Acehnese, and only four English vowels were comparable to Indonesian vowels. Acehnese had equivalent sounds for the English [iː], [ɛ], [ɔː], [uː], and [ʌ]. Indonesian had equivalent sounds for the English [iː], [ɛ], [ɔː], and [uː]. The semicolon-like symbols in front of some English vowels indicates that these vowels were produced for a longer duration compared to other English vowels. Both Acehnese and Indonesian vowels do not have tense and lax

features. The rest of the English vowels, [ɪ], [æ], [ɑː], and [ʊ] are novel to Acehnese and Indonesian vowel systems.

The equivalent of Acehnese and Indonesian vowels used to compare with English vowels are presented in Table 5.3. English vowels are classified into four categories, similar to Acehnese and Indonesian vowels (SAI), identical to Acehnese and Indonesian vowels (IAI), identical to Acehnese vowels (IA) and novel to Acehnese and Indonesian vowels (NAI). The first three sets of comparison comprise the English [iː], [ɔː] and [uː]. Acehnese and Indonesian vowels have similar vowels to these three English vowels but without tense features. This set was classified as similar to Acehnese and Indonesian (SAI). The second set contained only one English vowel [ɛ], which had an identical sound in Acehnese and Indonesian. This set is classified as identical to Acehnese and Indonesian (IAI). The third set also contained only one English vowel [ʌ] that had an identical sound only in Acehnese but not in Indonesian. This set was classified as identical to Acehnese (IA).

Table 5.3: Classification of English vowels to Acehnese and Indonesian vowels

Classification	English	Acehnese	Indonesian
SAI[1)]	[iː]	[i]	[i]
	[ɔː]	[ɔ]	[ɔ]
	[uː]	[u]	[u]
IAI[2)]	[ɛ]	[ɛ]	[ɛ]
IA[3)]	[ʌ]	[ʌ]	-
NAI[4)]	[ɪ]	[i]	[i]
	[ɑː]	[ɔ]	[ɔ]
	[ʊ]	[u]	[u]
	[æ]	[ɛ]	[ɛ]
		[a]	[a]

n.b. 1) Similar to Acehnese and Indonesian, 2) Identical to Acehnese and Indonesian, 3) Identical to Acehnese, 4) New to Acehnese and Indonesian

The last four sets comprised the English [ɪ], [æ], [ɑː], and [ʊ]. These four vowels are novel vowels for Ach-Ind speakers because the IPA symbol used to refer to these vowels is

absent from the Acehnese and Indonesian vowel systems. This set was classified as New to Acehnese and Indonesian (NA). In the previous section, it was found that the English [ɪ] and [ʊ] were conflated with the English [iː] and [uː], respectively. Thus, the Acehnese and Indonesian [i] are considered the equivalent vowels for the English [ɪ], while the Acehnese and Indonesian [u] were the equivalent vowels for the English [ʊ]. Based on an auditory examination and relative position in the vowel space, the English [ɑː] was found to be closer to the mid vowel of Acehnese and Indonesian [ɔ]. Thus, the Acehnese and Indonesian [ɔ] are considered equivalent vowels to the English [ɑː]. The impression of the English [æ] felt (upon hearing it) as being closer to the Acehnese and Indonesian [ɛ] but some speakers produced [æ] in the lower space closer to the Acehnese and Indonesian [a], so the Acehnese and Indonesian [a] were included in the analysis. In this comparison, only the F1 and F2 of similar vowels were compared. To determine whether the difference was significant, a statistical analysis was conducted to compare the F1 and F2 of the sounds.

5.4.1 Vowel [iː] in beat

The English [iː] is the first set in the SAI category. The scatter plot comparing the production of the English [iː], Indonesian [i] and Acehnese [i] by the Ach-IndE speakers are presented in Figure 5.3. It was apparent from the plot that the three vowels were congested in the upper left of the vowel space in both males and females. The three vowels were mostly conflated between 14 and 16 F2 bark in females and 13 and 15 in males. It was obvious that males and females produced vowels at slightly different locations. The male speakers' production of the three vowels tended to be higher and more to the back, while the females' production of the vowels tended to be lower and more to the front.

Another notable difference is that females' production of vowels was also more spread out while for the males it was accumulated in one area. The English [iː] and Indonesian

[i] were more spread out compared to the Acehnese [i] in female speakers. One male speaker also produced the Acehnese [i] at a more to the back location at 10 F2 bark, making it an outlier compared to another Acehnese [i]. Two female speakers also produced the English [i:] and Indonesian [i] at 13 F2 bark. Some female speakers also produced the English [i:] outside but still closer to the Acehnese [i] and Indonesian [i]. This visual inspection of the vowel spread suggests that the production of the three vowels was somehow conflated. However, a statistical test needed to be conducted to see if the three vowels were produced significantly different from each other.

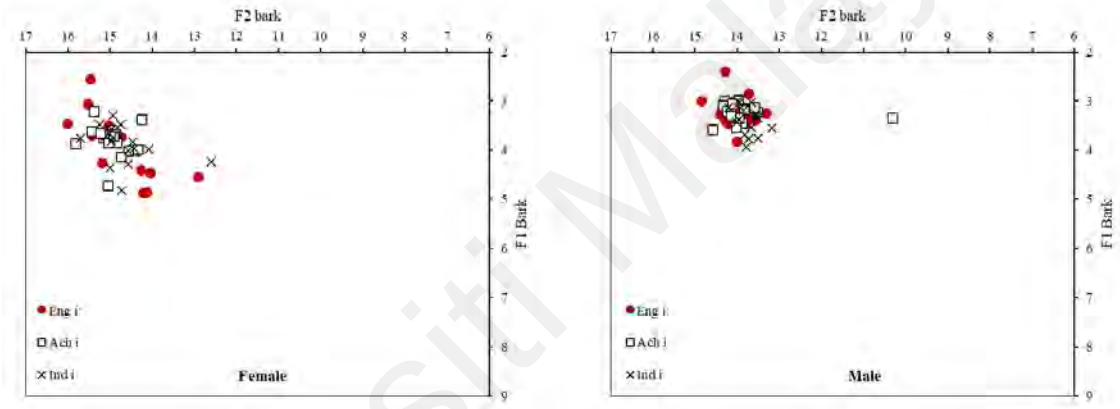

Figure 5.3: Plot of English [i:], Indonesian [i] and Acehnese [i]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [i:], Indonesian [i] and Acehnese [i]). For F1, Mauchly's Test of Sphericity was non-significant $p = 0.156$, so the result of Sphericity Assumed was used. The results indicated that the three vowels were not significantly different in terms of F1, $F(2, 56) = 2.360$, $p = 0.104$. As for F2, Mauchly's Test of Sphericity was non-significant $p = 0.427$, so the result of Sphericity Assumed was used. The result indicated that the three vowels were not significantly different in terms of F2 $F(2, 56) = 2.017$, $p = 0.143$. No post hoc test was conducted since there was no significant

difference for either F1 or F2 found among the three vowels. This finding confirmed the visual assumption that the English [iː], Indonesian [i] and Acehnese [i] were not distinguished in terms of quality. This result lent support to the L2LP theory that at the initial development, English learners tend to approximate non-native vowels to existing native vowels (Escudero, 2005).

5.4.2 Vowel [ɔː] in port

English [ɔː] was the second set in the SAI category. The scatter plot comparing the production of the English [ɔː], Indonesian [ɔ] and Acehnese [ɔ] by the Ach-IndE speakers are presented in Figure 5.4. The production of the English [ɔː] was mostly separated from the Indonesian and Acehnese [ɔ] in male speakers. The production of the English [ɔː] was also more spread out among females than males. The Acehnese and Indonesian [ɔ] produced by the males tended to accumulate in one area, while the two vowels were separated in the females. In female speakers, some instances of the English [ɔː] were produced in close proximity to the Acehnese and Indonesian [ɔ]. In both males and females, the English [ɔː] was produced higher and at a more to the back location. Acehnese and Indonesian [ɔ] were lower and more to the front, and they seemed to mix except for a few speakers who produced the Acehnese vowels in a more fronted position. A visual inspection of the vowel spread suggested that the production of the English [ɔː] differed from the Acehnese [ɔ] and Indonesian [ɔ]. However, a statistical test needed to be conducted to see if the three vowels were produced significantly different from each other.

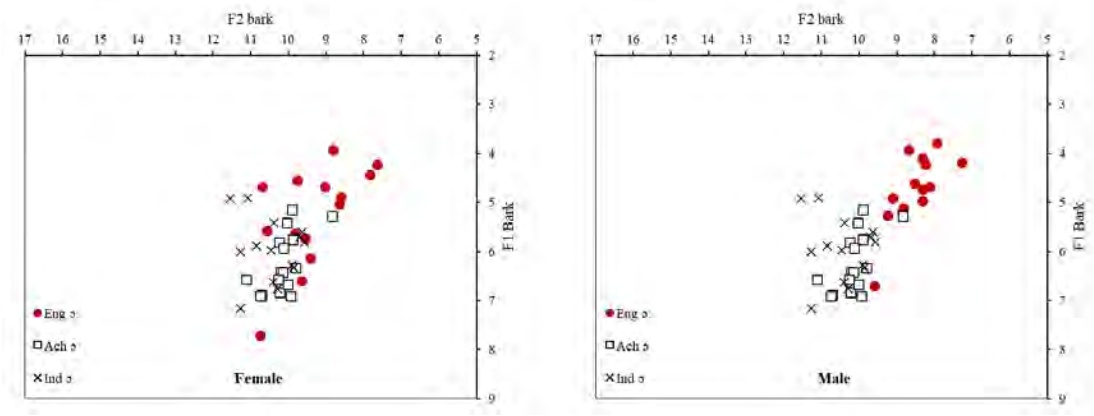

Figure 5.4: Plot of English [ɔː], Acehnese [ɔ], and Indonesian [ɔ]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [ɔː], Acehnese [ɔ], and Indonesian [ɔ]). For F1, Mauchly's Test of Sphericity was significant, the results of Hyun-Feldt were used with *epsilon* ε =0.84. The result indicated that the three vowels were significantly different in terms of F1, $F(2, 56) = 19.57$, $p < 0.001$. The difference was medium with partial η2 = 0.73, medium effect size. A post hoc comparison using a Bonferroni adjustment showed that F1 of the English [ɔː] was significantly different from the Acehnese [ɔ] ($p < 0.001$) with a large difference (Cohen's $d = 2.40$), and significantly different from the Indonesian [ɔ] ($p < 0.05$) with a large difference (Cohen's $d = 0.79$).

As for F2, Mauchly's Test of Sphericity was significant, so the results of Greenhouse-Geisser were used with *epsilon* ε =0.55. The results indicated that the three vowels were significantly different in terms of F2 $F(2, 56) = 11.68$, $p < 0.05$. The difference was medium with partial η2 = 0.80, medium effect size. A post hoc comparison using a Bonferroni adjustment showed that F2 of English [ɔː] was significantly different from the Acehnese [ɔ] ($p < 0.001$) with a large difference (Cohen's $d = 2.51$), and significantly different from the Indonesian [ɔ] ($p < 0.05$) with a large difference (Cohen's $d = 1.69$). This finding confirmed the visual assumption that the quality of the English [ɔː] was

different from the Indonesian [ɔ] and Acehnese [ɔ]. These findings may be attributed to the fact that the English [ɔ:] is produced with rhoticity pushing this vowel to the back of the vowel space.

5.4.3 Vowel [u:] in food

The English [u:] was the last set in the SAI category. The scatter plot comparing the production of the English [u:], Indonesian [u] and Acehnese [u] by the Ach-IndE speakers is presented in Figure 5.5. Apparently, the production of the English [u:] somehow separated from the Indonesian [u] and mixed with the Acehnese [u] to some extent. This pattern was mostly similar in both males and females. The Acehnese [u] was produced more fronted while the Indonesian [u] was produced more to the back. The production of the English [u:] sat in between the Acehnese and Indonesian [u]. Some instances of the English [u:] were clearly conflated with the Acehnese [u] and Indonesian [u], but the conflation was more pronounced toward the Acehnese [u]. A visual inspection of the vowel spread suggested that the production of the English [u:] differed from Indonesian [u] but not from Acehnese [ɔ] (especially for the males). Thus, a statistical test needed to be conducted to see if the three vowels were produced significantly different from each other.

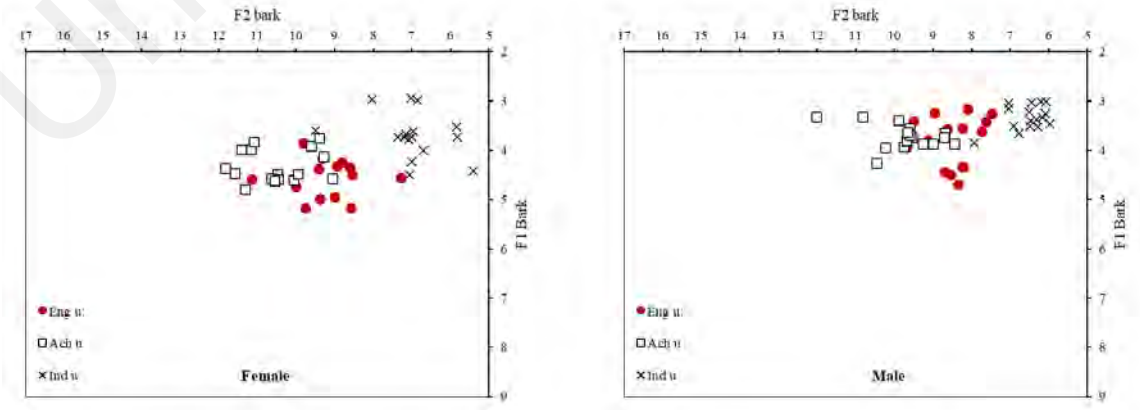

Figure 5.5: Plot of English [u:], Indonesian [u] and Acehnese [u]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [uː], Indonesian [u] and Acehnese [u]. For F1, Mauchly's Test of Sphericity was not significant, so the result of Sphericity Assumed were used. The results indicated that the three vowels were significantly different in terms of F1, $F(2, 56) = 37.917$, $p < 0.001$. The difference was small with partial $\eta 2 = 0.385$, a small effect size. A post hoc comparison using a Bonferroni adjustment showed that F1 of the English [uː] was not significantly different from the Acehnese [u] ($p = 0.286$) and the difference was large (Cohen's $d = 0.78$), but significantly different from the Indonesian [u] ($p < 0.001$).

As for F2, Mauchly's Test of Sphericity was also significant, so the result of Hunyh-Feldt was used with *epsilon* $\varepsilon = 1$. The results indicated that the three vowels were significantly different in terms of F2 $F(2, 56) = 149.32$, $p < 0.001$. The difference was small with partial $\eta 2 = 0.378$, a small effect size. A post hoc comparison using a Bonferroni adjustment showed that F2 of the English [uː] was significantly different from the Acehnese [u] ($p = 0.17$) and the Indonesian [u] ($p < 0.001$) and the difference is large (Cohen's $d = 1.25$).

This finding suggests that the quality of the English [uː] was a mix between the quality of the Acehnese [u] and the Indonesian [u]. In terms of F1, the English [uː] was closer to the Acehnese [u] while in terms of F2, the English [uː] was further from both Acehnese and Indonesian [u]. The speakers seemed to approximate the quality of the English [uː] through both Acehnese and Indonesian [u]. Compared to the Acehnese and Indonesian [i] and [ɔ], where their quality overlapped, the Acehnese and Indonesian [u] were separated. This separation might have played a role in squeezing the English [uː] in between the Acehnese and Indonesian [u].

5.4.4 Vowel [ɛ] in bet

The English [ɛ] was the only vowel classified as the IAI category. The scatter plot comparing the production of the English [ɛ], Indonesian [ɛ] and Acehnese [ɛ] by the Ach-IndE speakers is presented in Figure 2.6. Clearly, the production of the three vowels overlapped in both males and females. The conflation was more obvious in males than females. In males, the three vowels were crammed in one location as if they were not one indistinguishable vowel. For female speakers, the Indonesian [ɛ] was produced a little bit higher while the Acehnese [ɛ] was a little bit lower. Some instances of the English [ɛ] overlapped with the Indonesian [ɛ], while some overlapped with the Acehnese [ɛ]. A statistical test needed to be conducted to see if the three vowels were produced significantly different from each other.

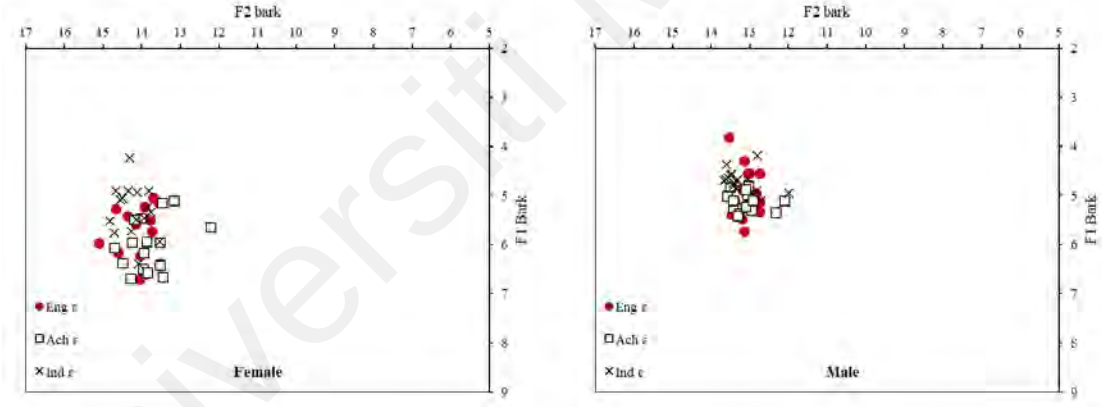

Figure 5.6: Plot of English [ɛ], Indonesian [ɛ] and Acehnese [ɛ]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [ɛ], Indonesian [ɛ] and Acehnese [ɛ]). For F1, Mauchly's Test of Sphericity was non-significant, so the result of Sphericity Assumed was used. The results indicated that the three vowels were significantly different in terms of F1, $F(2, 56) = 23.107$, $p < 0.001$. The difference was small with partial $\eta 2 = 0.452$, a small effect size. A post hoc comparison using a Bonferroni adjustment showed that F1 of the English [ɛ] was significantly different from the Indonesian [ɛ] ($p < 0.002$) and the

difference was medium (Cohen's d = 0.59), but not significantly different from the Acehnese [ɛ] (p = 0.060).

As for F2, Mauchly's Test of Sphericity non-significant, so the result of Sphericity Assumed was used. The result indicated that the three vowels were significantly different in terms of F2 $F(2, 56) = 4{,}910$, $p < 0.05$. The difference was small with partial $\eta 2 = 0.15$, a small effect size. A post hoc comparison using a Bonferroni adjustment showed that F2 of the English [ɛ] was not significantly different from the Acehnese [ɛ] (p = 0.082) and Indonesian [ɛ] (p = 0.066).

These findings suggests that the quality of the English [ɛ] is significantly different from the Indonesian [ɛ] in terms of F1 but not in terms of F2. The quality of the English [ɛ] and Acehnese [ɛ] were not different both in terms of F1 and F2. In other words, the English [ɛ] was fully conflated into the Acehnese [ɛ] and partially conflated into the Indonesian [ɛ].

5.4.5 Vowel [ʌ] in bud

The English [ʌ] was the only English vowel classified in the IA category in which only the Acehnese [ʌ] was considered identical to the English [ʌ]. However, since the location of the English [ʌ] produced by this speaker was assumed to be close to the Acehnese and Indonesian [a], these two vowels were included in the plot. The scatter plot comparing the production of the English [ʌ], Acehnese [ʌ], and the Acehnese and Indonesian [a] by the Ach-IndE speakers is presented in Figure 5.7. It can be seen that the quality of the English [ʌ] is different from the Acehnese [ʌ] in both males and females. Only two males and females conflated this vowel into the Acehnese [ʌ]. Most female speakers conflated the English [ʌ] into the Acehnese and Indonesian [a]. Two female speakers produced this vowel outside the location of the Acehnese and Indonesian [a]. One of the female speakers produced this vowel further in the front, and another female speaker produced it further

back. The Acehnese and Indonesian [a] were not totally conflated as in the Acehnese and Indonesian [i] and [ɔ].

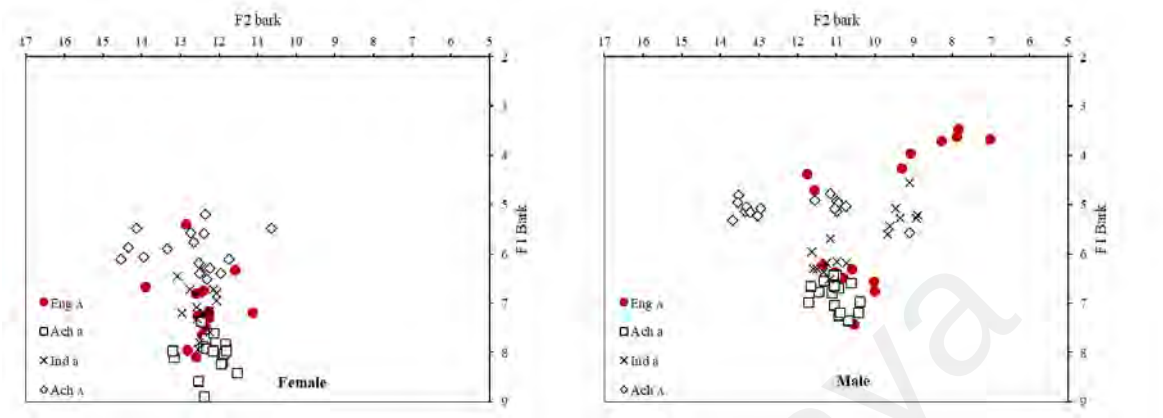

Figure 5.7: Plot of English [ʌ], Acehnese [ʌ], Indonesian [a] and Acehnese [a]

As for males, the vowel scatters in various location in the vowel space. Mostly they conflated the English [ʌ] into the Acehnese and Indonesia [a] at the bottom of the vowel space. Two male speakers approximated the English [ʌ] to the Acehnese [ʌ], but the rest of the speakers produced the English [ʌ] in the upper right of the vowel space. It was shown in section 1.2.2., that some male speakers produced the English [ʌ] closer to the English [uː] and [ʊ]. Since the English [uː] – [ʊ] are already conflated, it was assumed that some instances of the English [ʌ] would resemble the quality of the Acehnese and Indonesian [u]. Therefore, the English [ʌ] was projected against the Acehnese and Indonesian [u] as seen in Figure 5.8. Six male speakers (MS2, MS6, MS11, MS13, MS14 and MS15) approximated the English [ʌ] to the Acehnese and Indonesian [u].

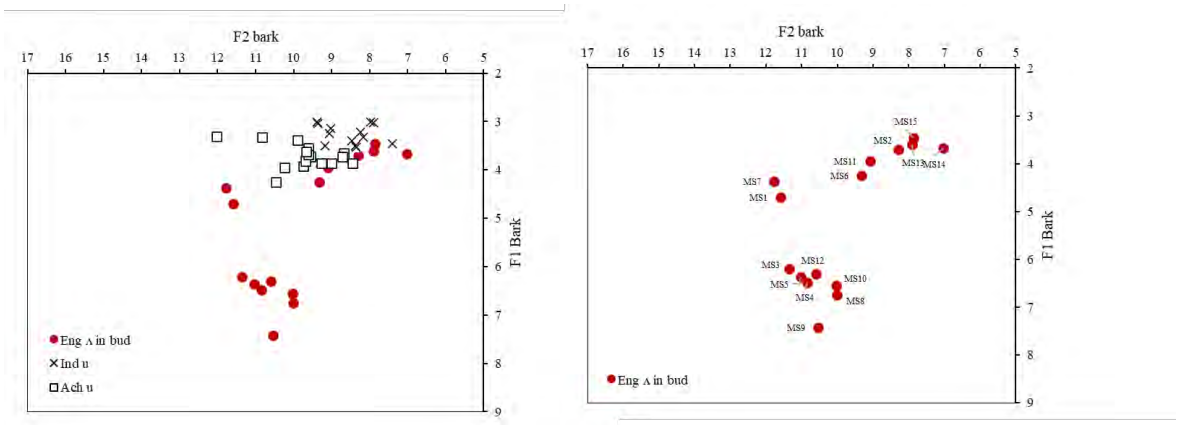

Figure 5.8: Plot of English [ʌ], Indonesian [u] and Acehnese [u]

The production of *bud* as [bud] instead of [bʌd] suggested that they relied on the spelling pronunciation of [u] in *bud*. The Indonesian letter *u* is produced as [u] in any Indonesian words such *buku* [buku] / *book*, *kutu* [kutu] / *lice*, and *lucu* [lutʃu] / *funny*. The same goes for the alphabet *i* as [i] and *a* as [a]. In Indonesian, allophones only occur in the words written in Indonesian alphabet e and o. The words containing Indonesian alphabet e can be pronounced as [ɛ] and [ə] while the alphabet o can be pronounced as [o] and [ɔ]. This assumption was further supported by the fact that most speakers were not familiar with the English *bud,* with an average familiarity index of 0.3. The familiarity index for all English words used in this study is presented in Table 5.4. Since they were not familiar with the word, it was likely that some speakers relied on Indonesian alphabetical sounds. Those who were not familiar with the word *bud* and approximated the pronunciation of *bud* to the Indonesian [a] might have been influenced by some English words such as *cut* [kʌt] and *but* [bʌt].

Table 5.4: Word familiarity index

Target Vowel	Target Words	Familiarity
[iː]	beat	0.3
[ɪ]	bit	0.5
[ɛ]	bet	0.1
[æ]	bat	0.6
[ʌ]	bud	0.3

[ɑ]	pot	0.6
[ɔ:]	port	0.6
[ʊ]	foot	1
[u:]	food	1
[ɝ]	bird	0.9

Since most speakers conflate the English [ʌ] to the Acehnese and Indonesian [a], a repeated-measures ANOVA was conducted. However, six speakers (MS1, MS2, MS6, MS7, MS11, MS13, MS14 and MS15) who did not approximate the English [ʌ] to the Acehnese and Indonesian [a] were excluded from the analysis.

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [ʌ], Indonesian [a] and Acehnese [a]). For F1, Mauchly's Test of Sphericity was significant, so the results of Greenhouse-Geisser were used with *epsilon* ε = 0.651. The results indicated that the three vowels were significantly different in terms of F1, $F(2, 56) = 26.158$, $p < 0.001$. The difference was small with partial η2 = 0.48, a small effect size. A post hoc comparison using a Bonferroni adjustment showed that F1 of English [ʌ] was significantly different from Acehnese [a] ($p < 0.001$) and the difference was large (Cohen's $d = 1.04$), but not significantly different from Indonesian [a] ($p = 0.435$).

As for F2, Mauchly's Test of Sphericity was significant, so the results of Hyundh-Feldt were used with *epsilon* ε = 0.78. The result indicated that the three vowels were significantly different in terms of F2 $F(2, 56) = 3.84$, $p = 0.038$. A post hoc comparison using a Bonferroni adjustment showed that F2 of English [ʌ] was not significantly different from Acehnese [a] ($p < 0.107$) and Indonesian [a] ($p = 0.18$). This finding suggests that the quality of the English [ʌ] was significantly different from the Acehnese [a] only in terms of F1 but not in terms of F2. The quality of the English [ʌ] and Indonesian [a] were not different both in terms of F1 and F2. In other words, the English

[ʌ] was fully conflated into the Indonesian [a] and partially conflated into the Acehnese [a].

5.4.6 Vowel [ɪ] in bit

The English [ɪ] was the first of four sets in the NAI category (Table 5.3). Different from the English [iː], the English [ɪ] is novel to Acehnese and Indonesian vowels. However, its close proximity to the Indonesian [i] and Acehnese [i] may encourage Acehnese-Indonesian speakers to produce the English [ɪ] around the native Acehnese and Indonesian [i]. The scatter plot comparing the production of the English [ɪ], Indonesian [i], and Acehnese [i] is presented in Figure 5.9. It was apparent from the plot that the three vowels were totally conflated in the vowel space for both males and females. The conflation was also more crammed in males than females. One outlier for the English [ɪ] and English [i] was present in males, and one outlier for the Indonesian [i] in females. The females' production of the three vowels seems to be separated between the upper left and lower right areas.

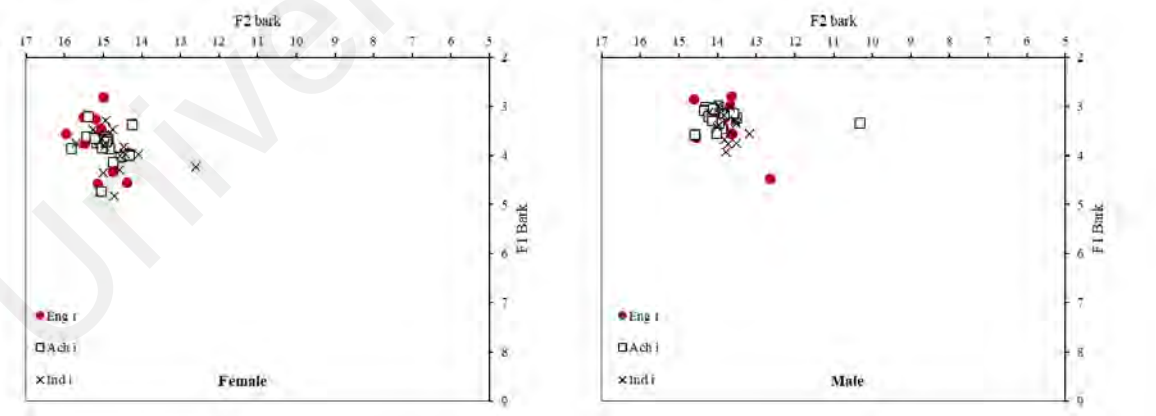

Figure 5.9: Plot of English [ɪ], Indonesian [i] and Acehnese [i]

Another notable finding was that the conflation of the English [ɪ] into the Indonesian [i] and Acehnese [i] was mostly similar to what was found in the English [iː] (section 2.4.1).

Since the English [iː] - [ɪ] were already conflated (section 1.3.1), the two vowels were also plotted against the Acehnese and Indonesian [i] in Figure 5.10 to better visualise the conflation of the four vowels. As expected, the four vowels were indistinguishable, especially in males. However, some female speakers slightly produced the English [iː] and [ɪ] outside the Acehnese and Indonesian [i]. Visual inspection of the vowel space suggested that the speakers did not discriminate between the four vowels. A statistical test was conducted to confirm this assumption if the English [ɪ] was produced significantly different from the Indonesian [i] and Acehnese [i].

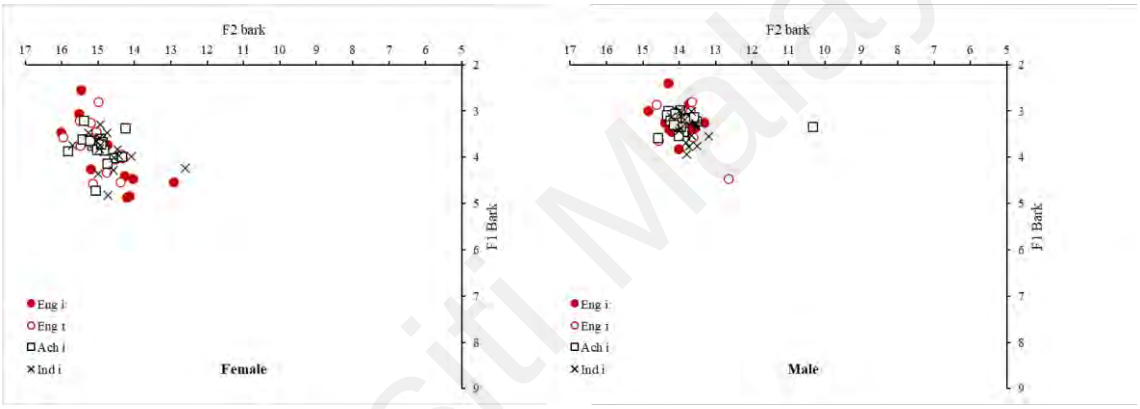

Figure 5.10: Plot of English [iː], English [ɪ], Acehnese [i] and Indonesian [i]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the four vowels from the three languages (i.e., English [iː], English [ɪ], Acehnese [i], and Indonesian [i]. For F1, Mauchly's Test of Sphericity was significant, so the results of Greenhouse-Geisser were used with *epsilon* ε =0.67. The results indicated that the four vowels were not significantly different in terms of F1, $F(3, 84) = 0.972$, $p = 0.385$. As for F2, Mauchly's Test of Sphericity was significant; the result of Hyun-Feldt was used with *epsilon* ε =0.87. The results indicated that the four vowels were not significantly different in terms of F2 $F(3, 84) = 1.37$, $p = 0.26$. No post hoc comparison was conducted for F1 and F2 since the findings were insignificant. These findings suggested that the English [iː], English [iː], Acehnese [i], and Indonesian [i] were both conflated in F1 and F2.

5.4.7 Vowel [æ] in bat

English [æ] is the second English vowel in the NAI category (Table 2.3). The Acehnese and Indonesian [ɛ] was the closest vowels to the English [æ]. However, since some speakers might have also produced *bat* based on its written orthography, the Acehnese and Indonesian [a] were included in the plot. The scatter plot comparing the production of the English [æ] Acehnese, Indonesian[ɛ] and Acehnese, and Indonesian [a] is presented in Figure 5.11. The male and female speakers seemed to be divided in producing the English [æ]. Most female speakers produced the English [æ] closer to Acehnese, and Indonesian [ɛ] and only a handful of them produced the English [æ] closer to the Acehnese and Indonesian [a]. On the other hand, male speakers produced the English [æ] around the Acehnese and Indonesian [ɛ] location, and some produced it around the Acehnese and Indonesian [a]. The graphical inspection seemed to suggest that the English [æ] was, to a certain degree, similar to the Acehnese and Indonesian [ɛ] and [a] vowels presented in the vowel space. A statistical analysis was conducted to confirm this assumption to compare the F1 and F2 of the English [æ] to the Acehnese and Indonesian vowels presented in Figure 5.11.

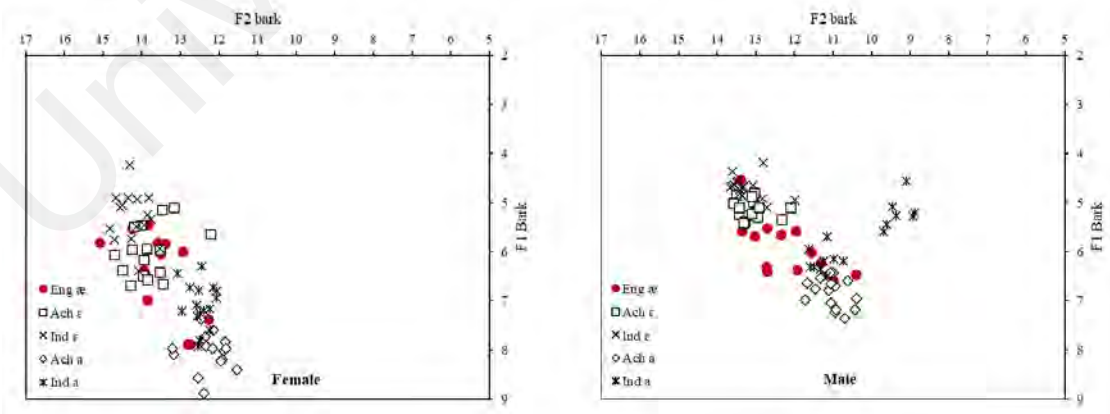

Figure 5.11: Plot of English [æ] Acehnese and Indonesia [ɛ] and Acehnese and Indonesian [a]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the five vowels from the three languages (i.e., English [æ], Acehnese [ɛ], Indonesian [ɛ], Acehnese [a] and Indonesian [a]). For F1, Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser was used with *epsilon* ε =0.69. The result indicated that the five vowels were significantly different in terms of F1, $F(4, 112) = 129.66$, $p < 0.001$. The difference was small with partial η2 = 0.822, a large effect size. A post hoc comparison using a Bonferroni adjustment showed that F1 of the English [æ] was significantly different from the Acehnese [ɛ] ($p < 0.002$), medium difference (Cohen's $d = 0.74$), the Indonesian [ɛ] ($p < 0.001$), large difference (Cohen's $d = 1.70$), the Acehnese [a] ($p < 0.001$), large difference (Cohen's $d = 1.8$), but not significantly different from the Acehnese [a] ($p = 0.484$).

As for F2, Mauchly's Test of Sphericity was significant, so the results of Greenhouse-Geisser were used with *epsilon* ε =0.63. The results indicated that the five vowels were significantly different in terms of F2 $F(4, 112) = 98.47$, $p < 0.001$. The difference was medium with partial η2 = 0.78, medium effect size. A post hoc comparison using a Bonferroni adjustment showed that F2 of the English [æ] was significantly different from the Acehnese [ɛ] ($p < 0.05$), medium difference (Cohen's $d = 0.67$), the Indonesian [ɛ] ($p < 0.001$), large difference (Cohen's $d = 0.9$), the Acehnese [a] ($p < 0.001$), large difference (Cohen's $d = 1.42$), and the Indonesian [a] ($p < 0.001$), large difference (Cohen's $d = 1.28$).

These findings suggest that the quality of the English [æ] was significantly different from the Acehnese [ɛ], Indonesian [ɛ], and Indonesian [a] except the Acehnese [a] in terms of F1 and significantly different from the Acehnese [ɛ], Indonesian [ɛ], Indonesian [a] and Acehnese [a] in terms of F2. In other words, the quality of the English [æ] was not

conflated into any Acehnese and Indonesian vowels tested except the Acehnese [a] in terms of F1.

5.4.8 Vowel [ɑː] in pot

The English [ɑː]is the third English vowel in the NAI category (Table 5.3). the Acehnese and Indonesian [a] are the closest vowels to the English [ɑː]. The scatter plot comparing the production of the English [ɑː], Indonesian [ɔ] and Acehnese [ɔ] is presented in Figure 5.12. It was clear that there was a lack of separation between English [ɑː], Indonesian [ɔ] and Acehnese [ɔ]. The conflation patterns between males and females were also different. In males, the conflation was more crowded due to the compactness of the three vowels. Few speakers produced the Indonesian [ɔ] slightly fronted, and one speaker also approximated the English [ɑː] around the fronted Indonesian [ɔ]. In females, the conflation was more stretched in terms of F1 from the bottom to the upper part of the vowel space. Based on the plot, it was assumed that the English [ɑː] differed from the Acehnese and Indonesian [a] but conflated with the Indonesian and Acehnese [ɔ]. A statistical test was conducted to confirm this initial assumption.

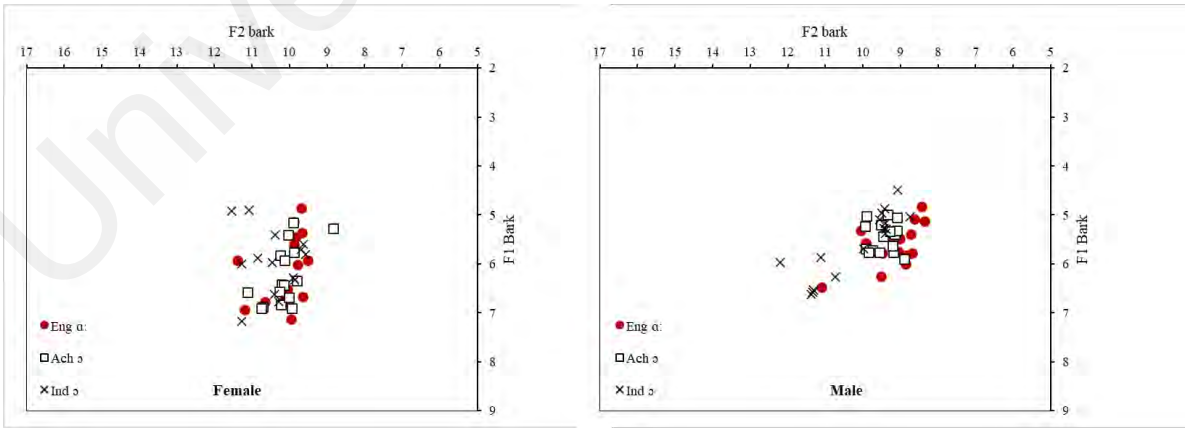

Figure 5.12: Plot of English [ɑː], Indonesian [ɔ] and Acehnese [ɔ]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [ɑː], Acehnese [ɔ], and Indonesian [ɔ]. For

F1, Mauchly's Test of Sphericity was non-significant, so the result of Sphericity Assumed were used. The results indicated that the three vowels were not significantly different in terms of F1, $F(2, 56) = 0.801$, $p < 0.454$. Since the results were not significant, no post hoc comparison was tested.

As for F2, Mauchly's Test of Sphericity was significant, so the results of Greenhouse-Geisser were used with *epsilon* ε =0.55. The results indicated that the three vowels were significantly different in terms of F2 $F(2, 56) = 5.492$, $p < 0.023$. The difference was small with partial η2 = 0.16, a small effect size. A post hoc comparison using a Bonferroni adjustment showed that F2 of the English [ɑː] was significantly different from the Indonesian [ɔ] ($p < 0.05$) with a small difference (Cohen's $d = 0.31$) but was not significantly different from the Acehnese [ɔ] ($p = 0.33$).

These findings suggest that the quality of the English [ɑː] was significantly different only from the Indonesian [ɔ] in terms of F2 but not in terms of F1. The quality of the English [ɑː] and Acehnese [ɔ] was not different both in terms of F1 and F2. In other words, the English [ɑː] was fully conflated to Acehnese [ɔ] and partially conflated to the Indonesian [ɔ].

5.4.9 Vowel [ʊ] in foot

The English [ʊ]is the last English vowel in the NAI category (Table 5.31). Different from the English [uː], the English [ʊ] is novel to Acehnese and Indonesian vowels. However, its close proximity to the Indonesian [u] and Acehnese [u] may encourage Acehnese-Indonesian speakers to produce the English [ʊ] around the two Acehnese and Indonesian back vowels. Thus, the Acehnese and Indonesian [u] are the closest vowels to the English [ʊ]. The scatter plot comparing the production of the English [ʊ], Indonesian [u] and Acehnese [u] is presented in Figure 5.13. It is obvious that the production of the three vowels was somehow conflated in the middle. The conflation was equally similar in male

and female speakers. The three vowels were similar in terms of F1 but differed across F2. The Acehnese [u] was produced to the left of the vowel space while the Indonesian [u] was produced to the right of the vowel space. Interestingly, the English [ʊ] was produced in between the Acehnese [u] and Indonesian [u], squeezing the two vowels away to make room for the novel vowel. As such, most instances of the English [ʊ] were conflated to the Acehnese [u] than the Indonesian [u]. The graphical analysis seemed to suggest that the English [ʊ] differed to some extent from the Indonesian [u] but not totally similar to the Acehnese [u]. To confirm this assumption, a statistical analysis was conducted to see if the three vowels were produced significantly different.

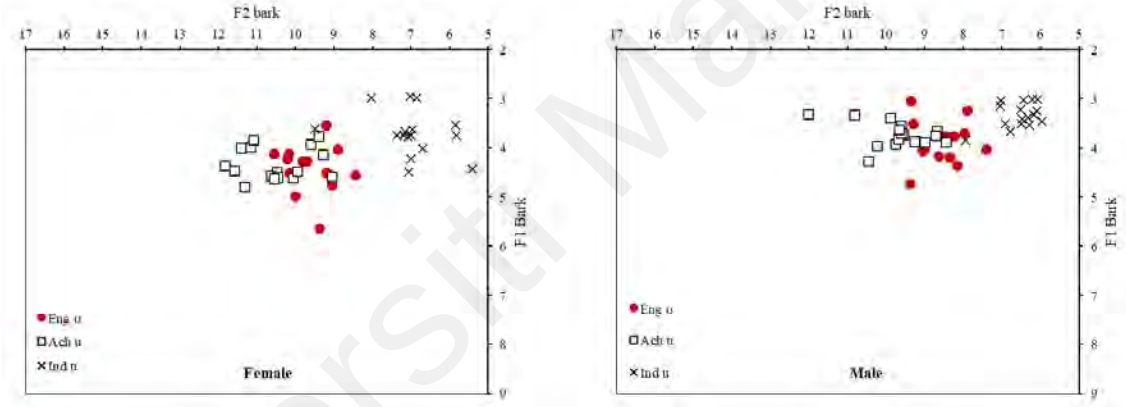

Figure 5.13: Plot of English [ʊ], Indonesian [u] and Acehnese [u]

A repeated-measures ANOVA was conducted to compare the F1 and F2 of the three vowels from the three languages (i.e., English [ʊ], Indonesian [u] and Acehnese [u]). For F1, Mauchly's Test of Sphericity was not significant, so the results of Sphericity Assmed were used. The results indicated that the three vowels were significantly different in terms of F1, $F(2, 56) = 38.90$, $p < 0.001$. The difference was moderate with partial $\eta 2 = 0.58$, a moderate effect size. A post hoc comparison using a Bonferroni adjustment showed that F1 of the English [ʊ] was not significantly different from the Acehnese [u] ($p = 0.48$) but significantly different from the Indonesian [u] ($p < 0.001$) and the difference was large (Cohen's $d = 1.19$).

As for F2, Mauchly's Test of Sphericity was also not significant, so the results of Sphericity Assumed were used. The result indicated that the three vowels were significantly different in terms of F2 $F(2, 56) = 183.02$, $p < 0.001$. The difference was large with partial $\eta 2 = 0.867$, a large effect size. A post hoc comparison using a Bonferroni adjustment showed that F2 of the English [ʊ] was significantly different from the Acehnese [u] ($p < 0.001$) and the Indonesian [u] ($p < 0.001$) and the difference was large (Cohen's $d = 1.84$). These findings suggested that the quality of the English [ʊ] was different from the Indonesian [u] in term of F1 and F2 but similar to the Acehnese [u] in term of F1.

Another notable finding is that the conflation of the English [ʊ] into the Indonesian [u] and the Acehnese [u] was mostly similar to what was found in the English [uː] (see section 2.4.3). Since the English [ʊ] - [u] were already conflated (see section 1.3.5), the two vowels were also plotted against the Acehnese and Indonesian [u] in Figure 5.14 to better visualise the conflation of the four vowels. It was clear that the English [ʊ] and [uː] were only conflated into the Acehnese [u] but not into the Indonesian [u]. One female speaker, however, produced the English [uː] closer to the Indonesian [u].

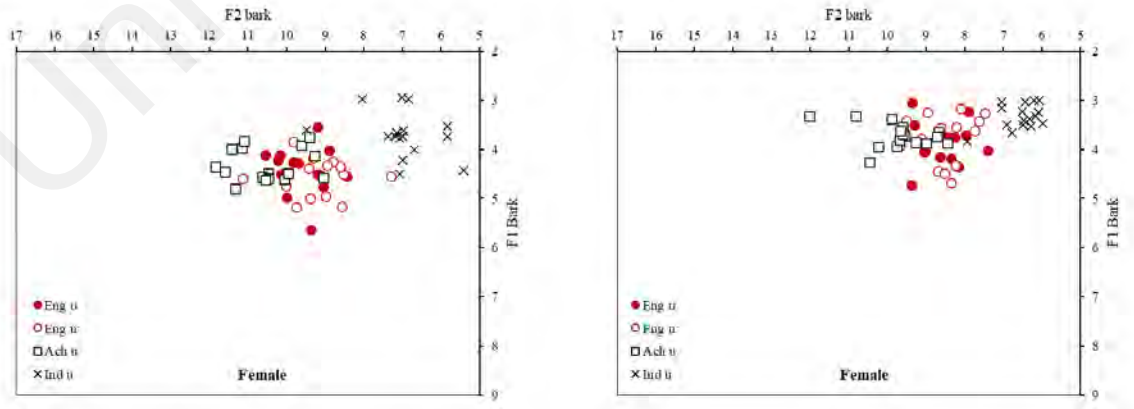

Figure 5.14: Plot of English [uː], Indonesian [u] and Acehnese [u]

5.5 Conclusion

This chapter compared the quality of English vowels to equivalent Acehnese and Indonesian vowels. The findings show that Ach-IndE speakers conflated English vowels into their existing L1 Acehnese and L2 Indonesian, and the conflation was more pronounced toward L1 Acehnese than L2 Indonesian. Based on the statistical analysis, there were five types of conflation. First, two English vowels were equally conflated into both L1 Acehnese and L2 Indonesian vowels: the English [ɪ] and English [iː], which were produced similarly to the Acehnese and Indonesian [i]. Second, one English vowel was produced similar to the equivalent Acehnese vowel but not the one in Indonesian. Third, one English vowel was fully conflated to L1 Acehnese but partially conflated into L2 Indonesian. Fourth, one English vowel was fully conflated into L2 Indonesian but partially conflated into Acehnese L1. Finally, one English vowel was partially conflated into both L1 Acehnese and L2 Indonesian. An example of this type of conflation is the English [uː] which was partially conflated into either the Acehnese [u] or Indonesian [u].

CHAPTER 6: PERCEPTION OF ENGLISH VOWELS

6.1 Introduction

This chapter presents the findings from the perception of English vowels by Acehnese-Indonesian speakers and discusses the relationship between the perception and production. First, the perception of five English vowel contrasts from three types of perception tasks (AX, ABX, FCI) were described individually. Second, the findings from the three tests were then compared to see if the speakers performed uniformly across the task types. Detailed analysis on the effect of the stimuli presentation order as also presented across the three tasks. Finally, the relationship between perception and production was explored by comparing it against the audio stimuli provided for the task by the American speaker. Perception and production of English vowels was also analysed using the Perceptual Assimilation Model (PAM) and Second Language Linguistic Perception (L2LP).

6.1 A comparison of English, Acehnese, and Indonesian vowels

Before going further into the perception analysis, the vowel systems were first compared across the three languages to determine which English vowel was new, similar, or identical to Acehnese and Indonesian vowels. The classification was made following the suggestions by Flege (1987). Escudero (2005) suggested that the L1/L2 comparison of vowel inventories could predict the ease of the acquisition of the L2 vowels. Best and Tyler (2007) also argued that listeners' relative ease and difficulty in perceiving non-native sounds could be predicted by comparing the phonetic similarities between L1 and L2. In determining similarity, the terms coined by Flege (1987) was used as in 'identical', 'similar', and 'new' to classify the L1/L2/L3 phonemic comparisons. Identical is then defined as sounds with similar IPA symbols in both languages and has similar acoustic features while similar have different acoustic features. The new scenario is when the two sounds are represented with different IPA symbols and different acoustic features. The

comparison of IPA symbols across the three languages (Acehnese, Indonesian and English) can be observed in Table 6.1.

Table 6.1: Vowel chart comparison across Acehnese (Asyik, 1987), Indonesian (Zanten, 1986) and American English (Ladefoged & Johnson, 2014)

English	Acehnese	Indonesian
[iː]	[i]	[i]
[ɪ]	[e]	[e]
[e]	[ɛ]	[ə]
[æ]	[ɯ]	[a]
[ʌ]	[ə]	[u]
[ɑː]	[ʌ]	[o]
[ɔː]	[a]	
[ʊ]	[u]	
[ɝ]	[o]	
[uː]	[ɔ]	

By comparing the IPA symbol and phonetic similarity across three languages, English vowel contrast could be classified into similar, new, or identical phonemes to Acehnese/Indonesia bilinguals. It is apparent from the Table 6.2 that only two English vowel pairs contain a similar vowel to Indonesian while there were two vowel pairs containing an identical vowel and three vowel pairs containing a similar vowel to Acehnese. English [iː] and [uː] had comparable sounds in both Acehnese and Indonesian and only differs in term of length. The English vowel [ʌ] and [ɔː] had equivalent sounds in Acehnese but the former was identical while the letter was similar in term of length. The English vowel [ɪ], [ʊ], [æ], [ɑː] and [ɑː] did not have comparable sounds in Acehnese and Indonesian and were classified as 'new'.

Table 6.2: Comparison of English vowel pairs (VP) to Acehnese and Indonesian based on the framework by Flege (1987)

	Identical-New	Similar-New	New-New
Indonesian	X	[iː] - [ɪ] (VP2) [uː] - [ʊ] (VP3)	[ɛ] – [æ] (VP1) [ʌ] – [ɑː] (VP4) [ɔː] - [ɑː] (VP5)
Acehnese	[ʌ] – [ɑː] (VP4) [ɛ] – [æ] (VP1)	[iː] – [ɪ] (VP2) [uː] – [ʊ] (VP3) [ɔː] – [ɑː] (VP5)	X

Based on this classification, the vowel pairs were then organized into four categories NNB (New-new to Acehnese-Indonesian), SNA (Similar-new to Acehnese), SNB (Similar-new to Acehnese-Indonesian) and INA (Identical-new to Acehnese). The classification can be observed in Table 6.2. NNB refers to English vowel pairs in which both vowels are new to Acehnese and new to the Indonesian vowel system. SNB refers to English vowel pairs in which one vowel was similar to the Acehnese and Indonesian vowel system and the other was new to the Acehnese and Indonesian vowel system. SNA refers to English vowel pairs in which one vowel was similar to the Acehnese vowel system and the other was new to the Acehnese vowel system. INA refers to English vowel pairs in which one vowel was identical to Acehnese and the other was new to the Acehnese vowel system. Note that, in the last two classifications, the comparison was only made to Acehnese. This was because Acehnese had two more vowels that did not exist in Indonesian vowel that was [ʌ] – [ɔː] based on Indonesian vowel system reported by Soderberg and Olson (2008) and Zanten and Heuven (1984). The classification of English vowel pairs could be observed in Table 6.3. This classification was used to interpret the perception performance for the AX, ABX and FCI tasks.

Table 6.3: Classification of English vowel pairs by NNB, SNA, SNB and INA categories

Category	Vowel Pairs
INA	[ʌ] – [ɑː] (VP4) [ɛ] – [æ] (VP1)
SNB	[iː - [ɪ] (VP2) [uː - [ʊ] (VP3)
SNA	[iː - [ɪ] (VP2) [uː - [ʊ] (VP3) [ɔː - [ɑː] (VP5)
NNB	[ɛ] – [æ] (VP1) [ʌ] – [ɑː] (VP4) [ɔː] – [ɑː] (VP5)

6.1.1 Findings from the AX Task

Figure 6.1 and Table 6.4 show the mean proportion of perception accuracy across five English vowel pairs [ɪ] – [iː], [æ] – [e], [ʊ] – [uː], [ʌ] – [ɑː], and [ɑː] – [ɔː] by Acehnese/Indonesian bilingual learners of English. The lowest proportion of accuracy was obtained for [æ] – [ɜː] while the highest accuracy was observed in [ʌ] – [ɑː]. Two vowel pairs, [æ] – [ɜː] and [ɑː] – [ɔː] sat at the bottom of the curve below 80 while the rest of the vowel pairs were above 95. The mean difference in proportion of accuracy between [æ] – [ɜː] and [ɑː] – [ɔː] was about 5 points at 75 and 80 respectively. While the mean accuracy for [ʌ] – [ɑː] almost reached a perfect score at 99, the values for vowel pairs [ʊ] – [uː] and [ɪ] – [iː] were slightly lower at 95 and 96 respectively. The mean difference between the vowel pairs with lower accuracy and higher accuracy seemed to be quite high at 15 points. An ANOVA test showed that there was a statistically significant difference in the perception accuracy across the five English vowel pairs tested ($p = 0.00 < 0.05$).

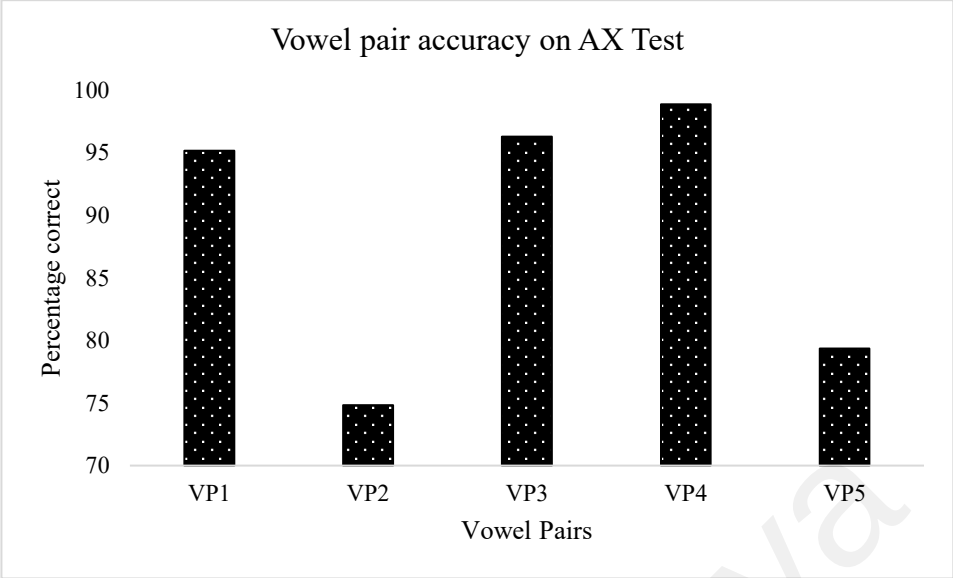

Figure 6.1: Mean percentage of accuracy of the AX task

A Tukey HSD post-hoc test (at a confidence level of 95%) was further conducted to see which of the specified vowel pairs differed from each other. The statistical analysis shows that the vowel pairs [æ] – [ɜː] and [ɑː] – [ɔː] were significantly different from the vowel pairs [ɪ] – [iː], [ʊ] – [uː] and [ɪ] – [iː] and vice versa. There was no difference in perception accuracy between vowel pairs [æ] – [ɜː] and [ɑː] – [ɔː] (VP2 × VP5, p = 0. 182) and no perception difference was observed between the vowel pairs [ɪ] – [iː], [ʊ] – [uː] and [ɪ] – [iː] (VP1 × VP3 p = 0.982, VP1 × VP4 p = 0.369, VP3 × VP4 p = 0.715). This indicates that the vowel pairs [æ] – [ɜː] and [ɑː] – [ɔː] were in the same subset homogenous group while the vowel pairs [ɪ] – [iː], [ʊ] – [uː] and [ɪ] – [iː] were in another subset homogenous group. The findings show that the speakers' perception may be priming to the 'new' vowel pairs (VP2) that are non-existent in both Acehnese and Indonesian and the pairs (VP5) in which one vowel exists only in Acehnese. The perception accuracy was higher in the vowel pairs (VP1 & VP3) in which one of the vowels similar to both Acehnese and Indonesian [iː] and [uː]. The perception accuracy was highest for the vowel pairs with one vowel identical only to Acehnese (VP4).

Table 6.4: Mean score and Standard Deviation for each English Vowel Pairs

	Vowel Pairs	Mean	SD
VP1	ɪ - iː	95.16	9.35
VP2	e - æ	74.84	8.80
VP3	ʊ - uː	96.29	7.07
VP4	ʌ - ɑː	98.87	2.80
VP5	ɑː - ɔː	79.35	10.06

In order to see how the degree of similarity between L1/L2 contributes to perception accuracy of English vowel pairs, the data from Table 10 was plotted into similarity categories as presented in Figure 6.2. The figure clearly show that the speakers scored the lowest accuracy for the English vowel pairs that do not exist in Acehnese and Indonesian (NNB) and this is followed by the English vowel pairs in which one vowel is similar to Acehnese while the other new to Acehnese and Indonesian (SNA). The vowel pairs in which one vowel is similar to Acehnese and the other is new to Acehnese and Indonesian (INA) reached the highest accuracy at 99. The speakers seem to consistently perceive English vowel pairs in which one of the pairs exist in Acehnese and Indonesian at a near similar level, that is 95 and 96 respectively.

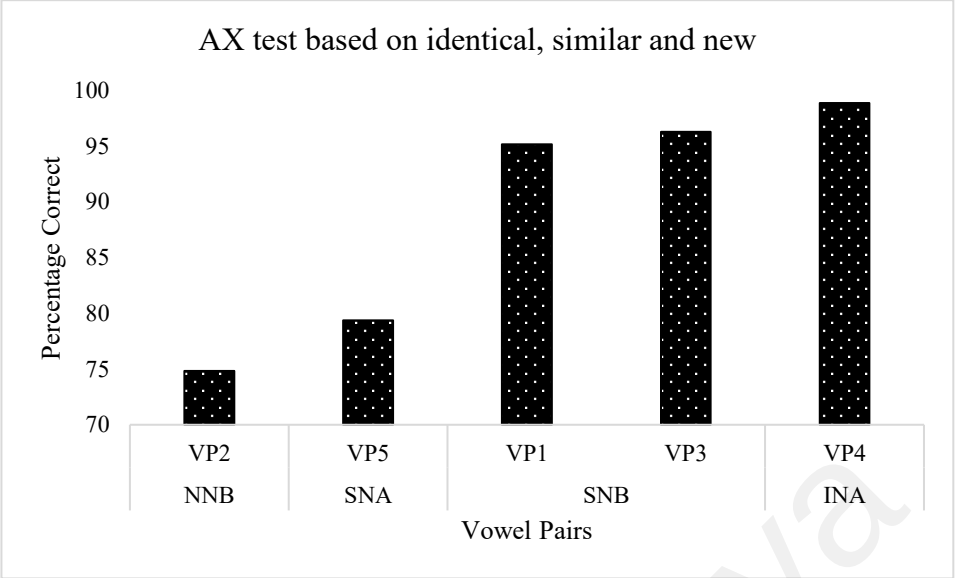

Figure 6.2: Mean percentage of accuracy based on NNB, SNA, SNB, and INA categories

In order to see if there is any perceived difference in terms of trial types in each vowel pairs, the data were further divided based on the order of presentation AA, AB, BA and BB. Figure 6.3 depicts the mean proportion of correct answers based on trial types (repetition and minimal pairs) in AX task across five English vowel pairs. Overall, the Ach-IndE speakers managed to consistently scored above 80 for every trail type across the vowel pairs except for BA and AA sequences for the vowel pairs [æ] – [ɜː] and [ɪ] – [iː] respectively. For AA sequence in VP3, speakers only correctly identified repetition at around 25%. For BB sequence in VP5, speakers only correctly identified repetition at slightly repetition 30%. For three other pairs, trial type order (minimal or repetition) did not seem to influence their score.

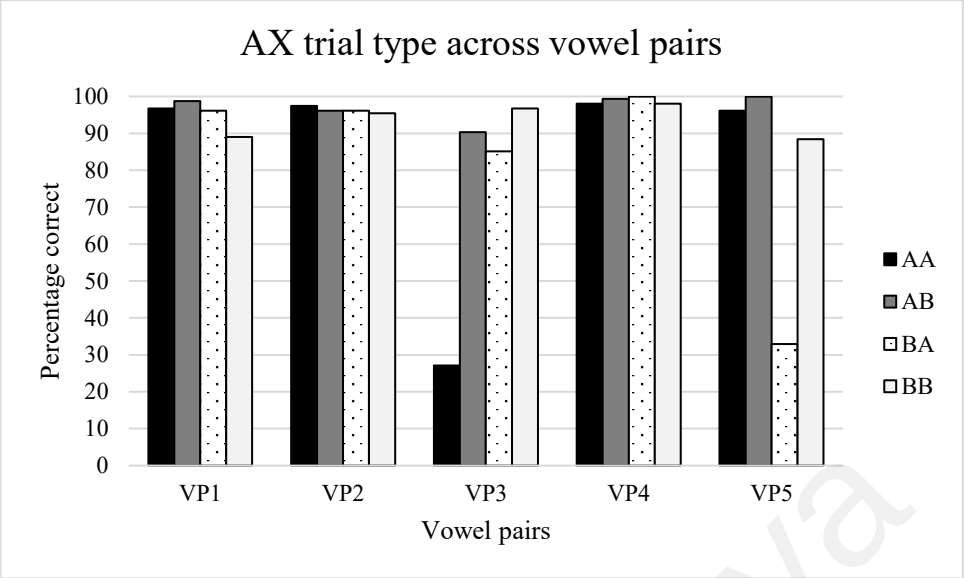

Figure 6.3: AX Mean percentage of accuracy based on trial type

Figure 6.4 compares the vowel perception accuracy in the AX task divided based on vocabulary size categories. Overall, the three groups of speakers performed equally comparable in all five vowel pairs. However, the difference was observed when looking at individual pairs where each group performed disproportionately within each pair. The high performing group perform a little bit higher in VP1 but slightly lower for VP2 and VP3. The high performing group scored 73 and 75 points for VP2 and VP5 while the figure for the rest of the pair was above 90 points. In VP4, the three groups had a similar score at around 98 points. On the other hand, the low performing speakers had the highest score compared to the other two groups in VP5 at 81 points. The perception accuracy for the medium group was the highest in VP3 at 98 points but lowest in VP5 at 75 points. The medium performing speakers had highest score in VP3 at 98 points while the high performing and low performing group had closely similar scores at 93 and 94 points respectively. Thus, it can be said that speakers' perception accuracy on the AX task did not appear to be influenced by their vocabulary size score.

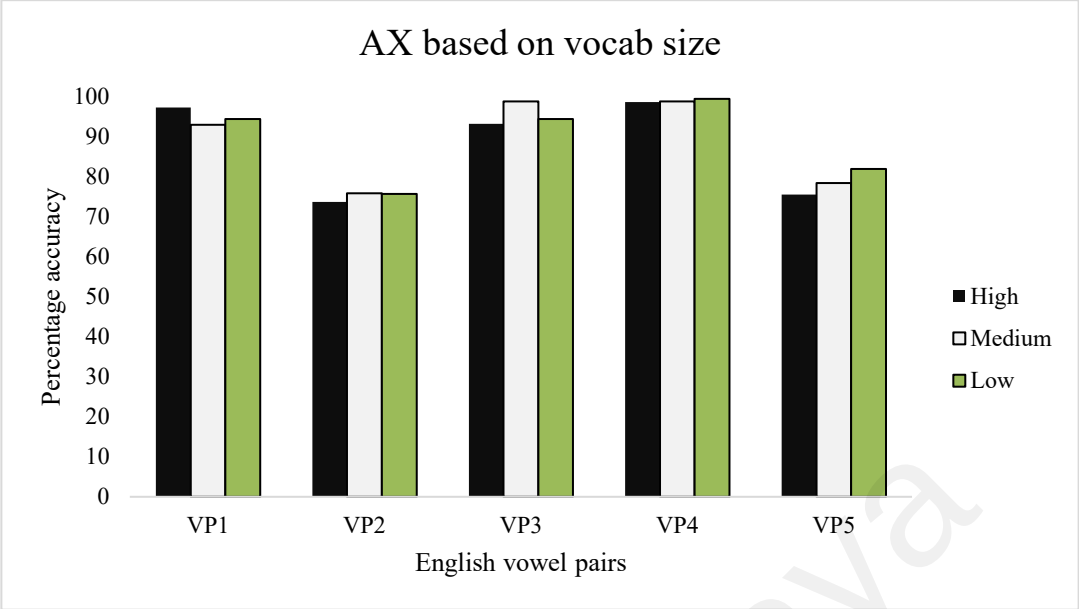

Figure 6.4: AX perception accuracy based on vocab size categories

Figure 6.5 compares perception accuracy to speakers' lexical knowledge of the words used in the study. Lexical knowledge did not seem to affect their perception of vowel pairs in the study. The students' perception accuracy on vowel pairs containing words that were less familiar to the speakers was similar to vowel pairs containing words they were familiar with. For example, VP4 had words that were less familiar than VP5 and VP3, and yet their perception accuracy was similar for these three pairs.

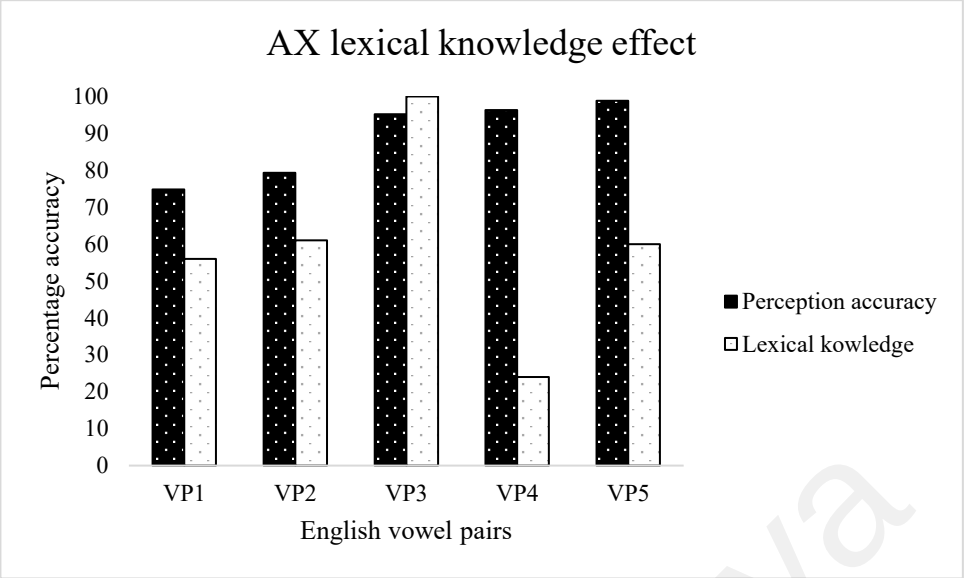

Figure 6.5: Comparison of AX perception accuracy and lexical knowledge

6.1.2 Findings from ABX Task

This section discussed the second task used to collect perception data, the ABX task. The task was administered immediately after they completed the AX task. In this task, speakers listened to three sounds instead of two in the AX task. They had to determine whether the last sound resemble the first and the second sound. The findings of the perception task can be observed in Figure 6.6.

Figure 6.6. shows the mean proportion of perception accuracy across five English vowel pairs [ɪ] – [iː] (VP1), [æ] – [ɜː] (VP2), [ʊ] – [uː] (VP3), [ʌ] – [ɑː] (VP4), and [ɑː] – [ɔː] (VP5) by Acehnese/Indonesian bilingual learners of English in the ABX task. The speakers appeared to perform better in the ABX task compared to the AX task. Speakers could attain a score of above 90 percent for all the vowel pairs in the ABX task. The lowest proportion of accuracy was obtained for VP1 while the highest accuracy was observed forVP2 and VP4. The mean accuracy for the highest vowel pairs almost reached a perfect score at 99 percent while the pairs with the lowest accuracy, VP1 was at 90

percent. The mean difference in proportion of accuracy between VP2 and VP4 is about 5 points at 93.39 and 97.42 respectively which suggests that the speakers were equally good at discriminating the vowels between these two pairs.

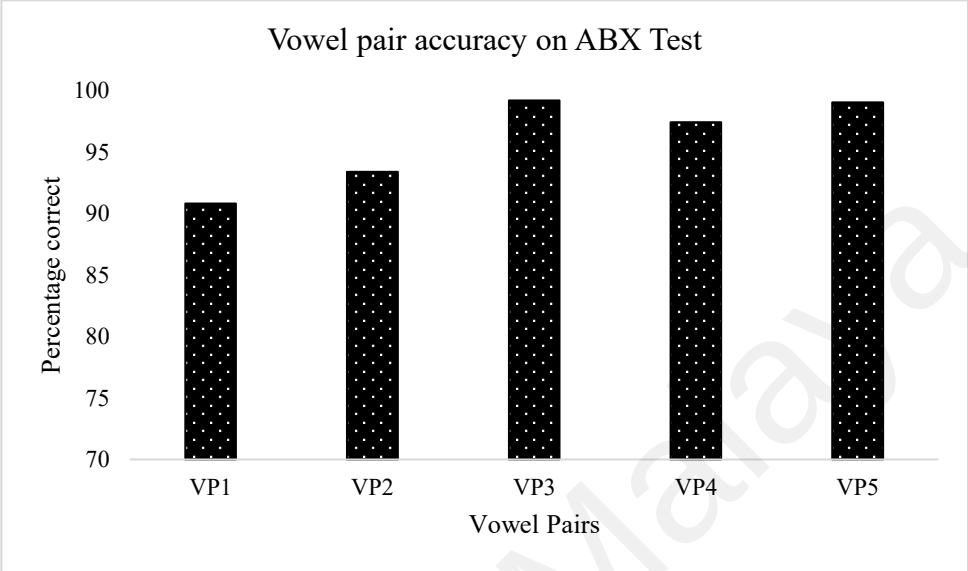

Figure 6.6: Mean percentage of accuracy of ABX task

A detailed comparison of mean average and standard deviation can be observed in Table 6.5. This suggests that there was a large variability in terms of speakers' performance for these pairs. The larger standard deviation for this pair also contributed to the lower performance of accuracy compared to other pairs. The standard deviation for VP2 and VP4 was at 7.35 and 6.04 respectively. Their performance was equal for the highest pairs VP3 and VP5 at 2.27 and 2.39 standard deviation.

Table 6.5: Mean score and Standard Deviation for ABX task

	Vowel Pairs	Mean	SD
VP1	[ɪ] – [iː]	90.81	11.04
VP2	[e] – [ɜː]	93.39	7.35
VP3	[ʊ] - [uː]	99.19	2.27
VP4	[ʌ] – [ɑː]	97.42	6.04
VP5	[ɑː] - [ɔː]	99.03	2.39

Similar to the AX task, to see how the degree of similarity between L1/L2 contributed to perception accuracy of English vowel pairs, the data from Table 6.5 was plotted into similarity categories (see Figure 6.7). VP1 and VP4 contained one vowel identical to Acehnese and was classified as INA. VP2, VP3, and VP5 contained one vowel similar to Acehnese and was classified as SNA. SNB is for VP2 and VP3 in which one as was similar to Indonesian. NNB contained one vowel which was new to Indonesian. However, the pattern did not indicate that this classification provided any meaningful contribution to the perception accuracy of the speakers. The SNB had pairs with lowest score and highest score at the same time. On the other hand, VP2 had a lower mean score than the other pairs in the SNA group and its mean score was lower than VP1 which contained identical vowels. However, it was important to note that for the ABX task, the speakers' performance in all pairs were already above 90 percent. Factors other than first language influence seemed to play a role in shaping the speakers' perception accuracy.

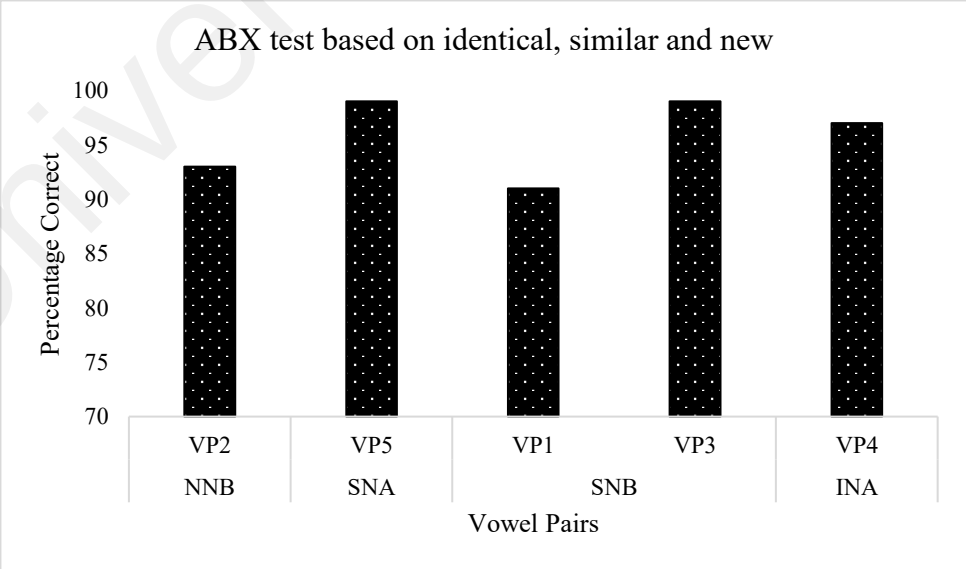

Figure 6.7: Mean percentage of accuracy based on NNB, SNA, SNB, and INA categories for ABX

In order to see if there was any perceived difference in terms of trial types in each vowel pairs, the data was presented based on the order of presentation AA, AB, BA and BB. Figure 6.8 depicted the mean proportion of correct answers based on trial types (repetition and minimal pairs) in the ABX task across five English vowel pairs. Overall, the Acehnese/Indonesian bilingual learners of English managed to consistently score above 90 for every trail type across the vowel pairs except for ABB sequences for the VP3. For VP3, the score for ABB sequence was slightly below 80%. This suggests that their performance across the vowel pairs was consistently higher regardless of the trial type orders.

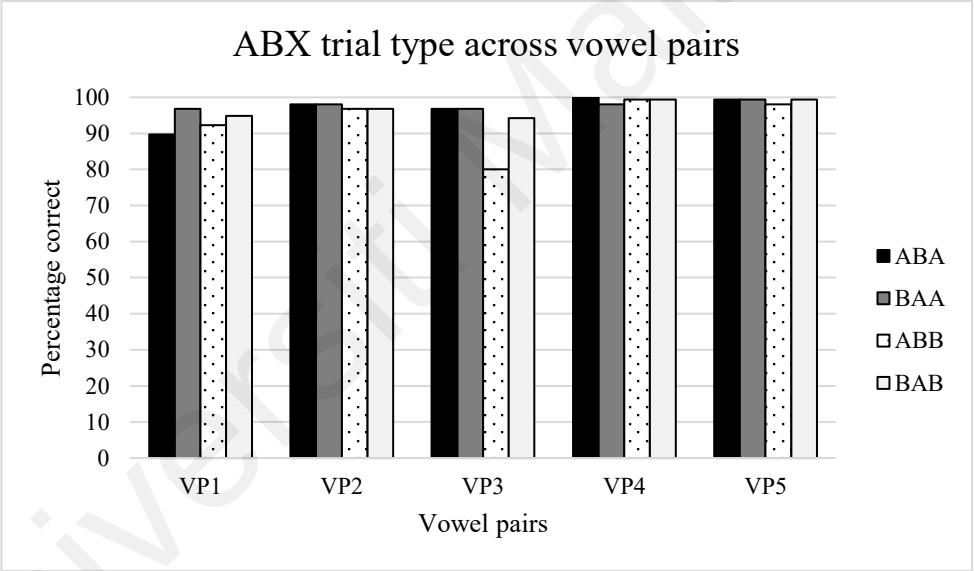

Figure 6.8: ABX Mean percentage of accuracy based on trial type.

Figure 6.9 compares the vowel perception accuracy in AX task divided based on the vocabulary size categories. The three groups of speakers performed equally in all five vowel pairs. The high-performing group scored above 91 points for VP3, VP4 and VP5 but the figure slightly lower for VP1 and VP2 at 90 and 92 points respectively. The perception accuracy for the medium group was, on the other hand, higher in VP3, VP4 and VP5 compared to other groups. The medium-performing speakers gained a perfect score for VP3 at 99 points and this was higher than that of the higher-performing group.

Similar pattern was found in VP1 and VP2. The medium performing group scored higher for this pair compared to the higher-performing group. The score obtained by the higher-performing group was also lower than the higher-performing group. The low-performing group had consistently lower scores than the higher-performing group for all pairs except for VP2 which was just three points above the high-performing group.

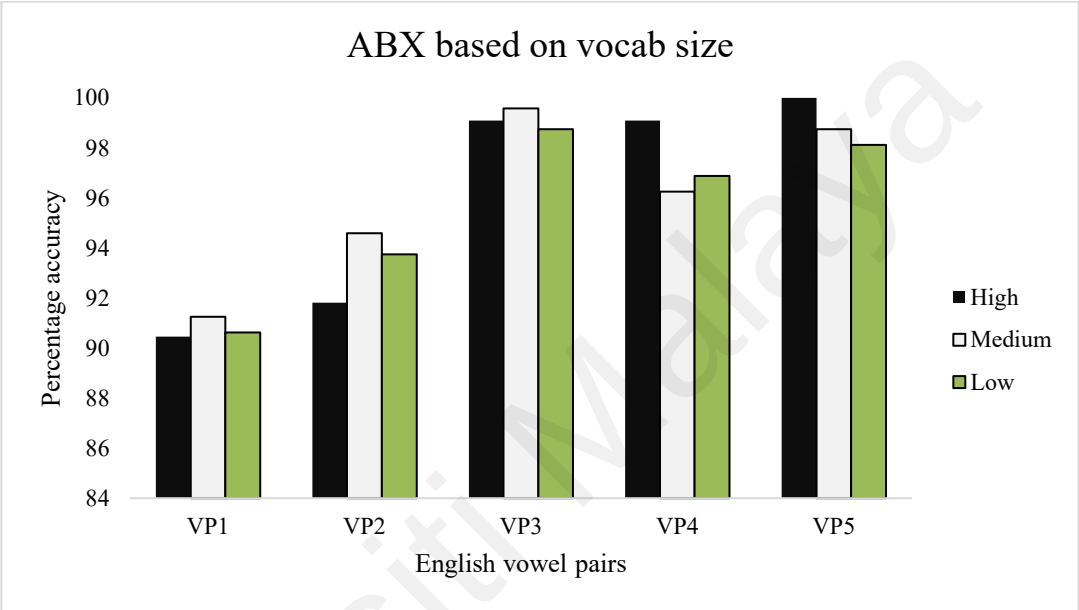

Figure 6.9: ABX perception accuracy based on vocab size categories

Figure 6.10 compares perception accuracy in the ABX task to speakers' lexical knowledge of the words used in the study. Lexical knowledge did not appear to affect their perception of vowel pairs in the study. The speakers performed the highest in VP3 and VP5 despite the fact that around 40% of the speakers were not familiar with one or any word in VP5 but were familiar with the two words in VP3. However, it was important to note that the American speaker had a rhotic accent, and thus produced the post-vocalic *r* in his production of the word *port*. Thus, it was highly likely that the speakers relied on rhoticity instead of their familiarity of the word. VP4, in which speakers were not familiar with the words in the pairs had a perception accuracy of 97. Speakers had lower lexical knowledge for both VP1 and VP2 , and yet, they managed to score higher in this task.

Lexical accuracy was, therefore, also not a good predictor for their perception accuracy performance in the vowel pairs tested.

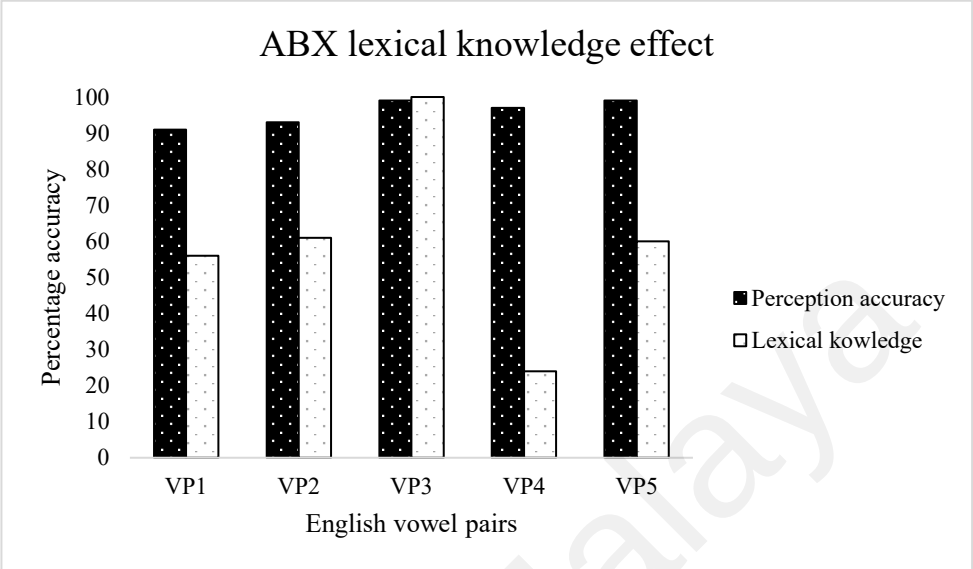

Figure 6.10: Comparison of ABX perception accuracy and lexical knowledge

6.1.3 Findings from Forced Choice Identification (FCI) Task

Figure 6.11. shows the mean proportion of perception accuracy across five English vowel pairs [ɪ] – [iː] (VP1), [æ] – [ɜː] (VP2), [ʊ] – [uː] (VP3), [ʌ] – [ɑː] (VP4), and [ɑː] – [ɔː] (VP5) by Ach-IndoE speakers in the forced choice identification task. The perception accuracy for the FCI task was overall lower than the accuracy in the previous task. This was due to the nature of the task in which the speakers need to choose which of the two words provided resemble the sound they listen to. VP2 had the lowest perception accuracy while VP3 had the highest perception accuracy. Both VP2 and VP3 had one vowel similar to Acehnese and Indonesian vowels. VP3 contained one vowel similar only to Acehnese and not to Indonesian. The perception accuracy for this vowel pair was the second lowest in the pairs. VP1 and VP4 which contained one vowel new to Indonesian but which is identical to Acehnese also had an uneven perception accuracy. VP1 had a

lower perception accuracy compared to VP4. Mean score of SD for each English pair is presented in Table 6.6.

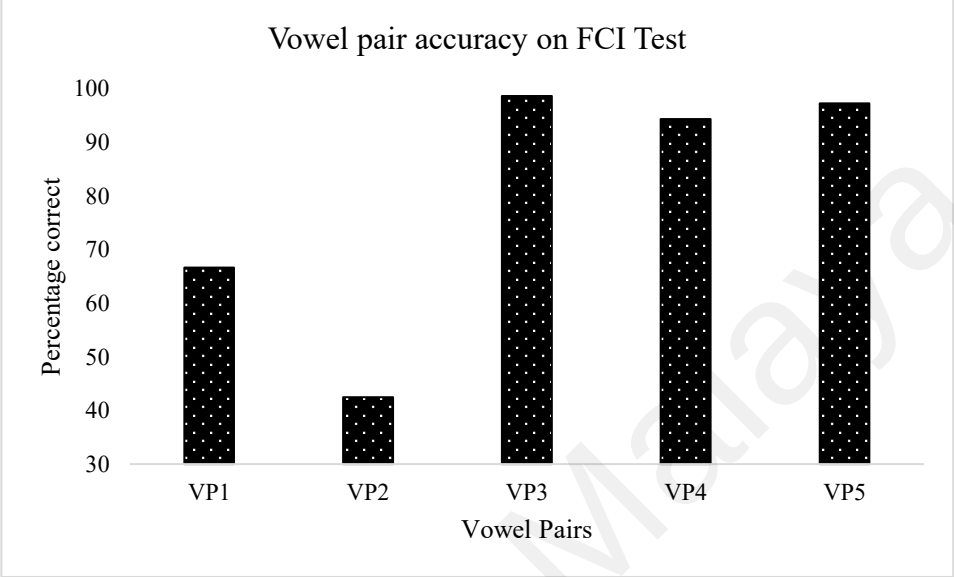

Figure 6.11: Mean percentage of accuracy of FCI task

Table 6.6: Mean score and Standard Deviation for each English Vowel Pairs

	Vowel Pairs	Mean	SD
VP1	[ɪ] – [iː]	66.67	40.54
VP2	[e] – [æ]	42.47	39.17
VP3	[ʊ] – [uː]	98.66	4.36
VP4	[ʌ] - [ɑː]	94.35	11.26
VP5	[ɑː] – [ɔː]	97.31	8.44

The classification of vowel pairs into INA (identical and new to Acehnese), SNA (similar and new to Acehnese), SNB (similar and new to Indonesian), and NNB (new and new to Indonesian) can be observed in Figure 6.12. The speakers scored the highest for the SNA and INA pair. The NNB pair had the lowest score for the FCI task. The result for SNB is uneven for VP1 and VP3.

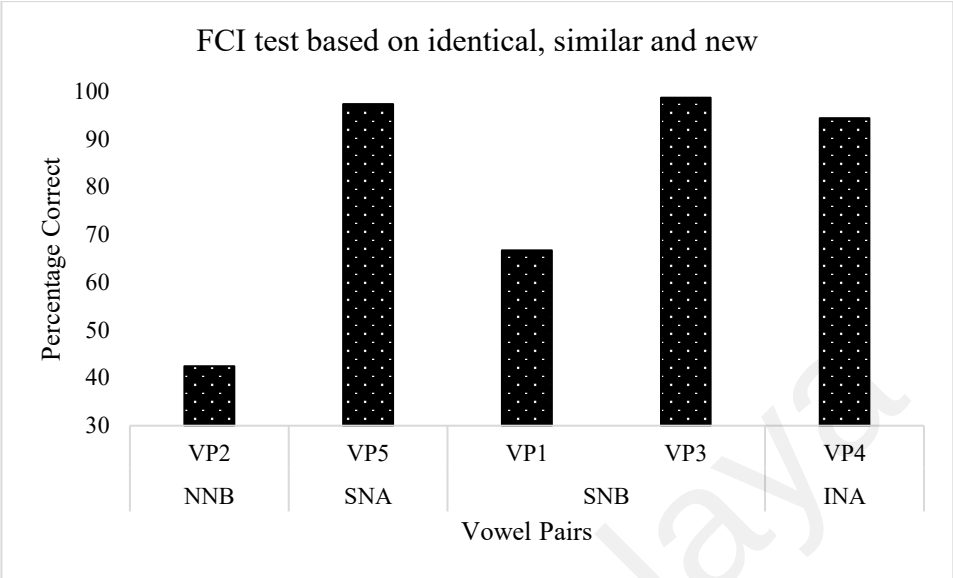

Figure 6.12: Mean percentage of accuracy based on NNB, SNA, SNB, and INA categories for FCI

Figure 6.13 depicts the mean proportion of correct answers based on trial types (repetition and minimal pairs) in the FCI task across five English vowel pairs. Overall, the Ach-IndE speakers of English managed to consistently discriminate across the trial type. They might have had a certain perception of a sound locked to the target sound and thus, made a decision based on that locked perception. However, the speakers tended to identify B better than A in VP1 and VP4. Their perception ability to identify the vowel was uniform in the rest of the pairs. Examples of trial type orders can be observed in Table 6.7.

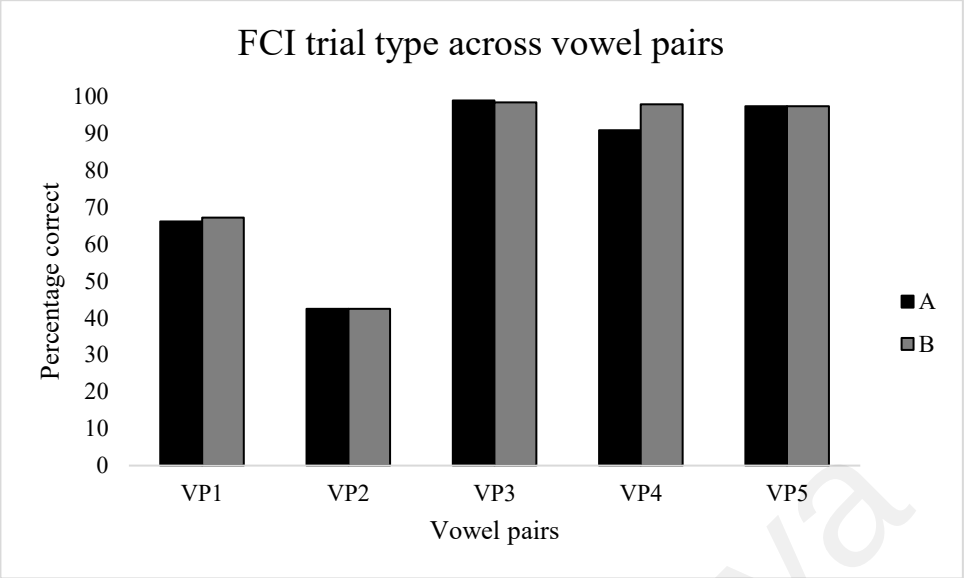

Figure 6.13: FCI Mean percentage of accuracy based on trial type.

Table 6.7: FCI trial type orders for [ɪ] – [i:]

Vowel Pair	Order	Stimuli	Repetition	Total
[ɪ] – [i:] (VP1)	ABB	beat - bit - bit	five rep.	20
	ABA	beat - bit - beat	five rep.	
	BAA	bit - beat - beat	five rep.	
	BAB	bit - beat - bit	five rep.	

Figure 6.14 compares the vowel perception accuracy in FC task divided based on vocab size categories. Black bar indicated student with higher score in vocabulary size test, followed by white bar for medium and green for low group. The three groups of speakers performed equally comparable in all five vowel pairs. The high performing group consistently perform a little bit higher in vowel pairs. The medium group had a slightly better perception in VP1, VP2 and VP3 while the perception accuracy for the VP4 and VP5 was equally comparable for the medium and low group.

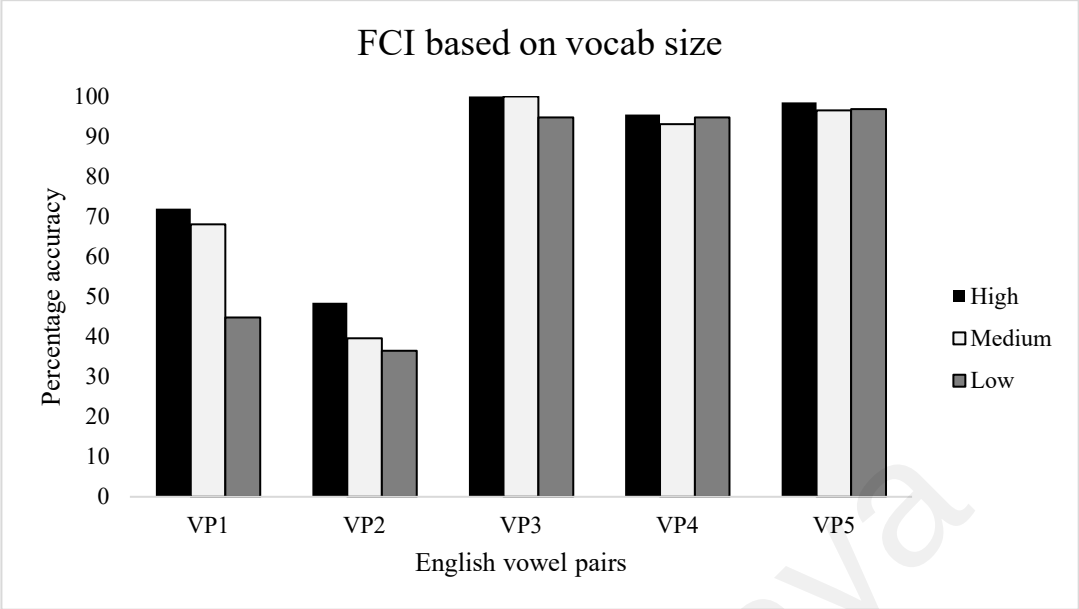

Figure 6.14: Comparison of FCI perception score and vocab size categories

Figure 6.15 compares the perception accuracy in the FCI task to speakers' lexical knowledge of the words used in the study. Similar to the findings in the previous two tasks, lexical knowledge was not a good prediction for the perception accuracy. The students' performance was the same regardless of their lexical knowledge of the words tested. The speakers performed the highest in VP3 and VP5 despite the fact that 50% of the speakers were not familiar with one or any of the words in the VP5 but familiar with all the words in VP3. VP4, in which speakers had lowest lexical knowledge had a perception accuracy of 97. Speakers had lower lexical knowledge for VP1 and VP2 and their perception accuracy was also lower. A regression analysis was further conducted to see if lexical knowledge had anything to do with their perception accuracy in the FCI task.

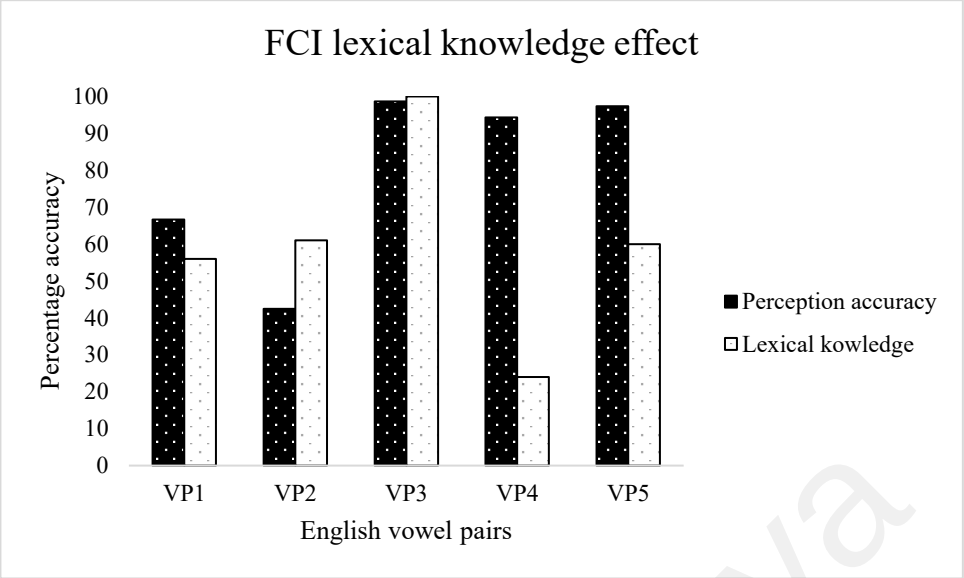

Figure 6.15: Comparison of FCI perception accuracy and lexical knowledge

6.1.4 Comparisons of the Three Tasks

In order to see how speakers, perform perception accuracy across the three different tasks, the mean average of the five vowel pairs was projected into a line graph. First, the data for all the speakers across three tasks are compared. Then, the data based on the vocabulary size categories was presented to see how speakers with different proficiency perform across the three tasks. ANOVA repeated measures test was conducted to see if different tasks yielded significantly different findings for each vowel pair.

Figure 6.16 compares the perception accuracy of five English vowel pairs in three different tasks, AX, ABX and FCI by Acehnese-Indonesian multilingual speakers. Overall, it was clear that VP2 had the lowest perception accuracy across the three tasks. VP2 was the most difficult pair in the FCI task. VP2 and VP5 are the most difficult pairs in the AX task. For the ABX task, it is VP1 that has lower perception accuracy. VP3 and VP4 sat at the upper graph in all three tasks indicating its difficulty to discriminate was uniform across the three tasks. A detailed comparison of average perception accuracy for each vowel pair across the three tasks can be observed in Table 6.14.

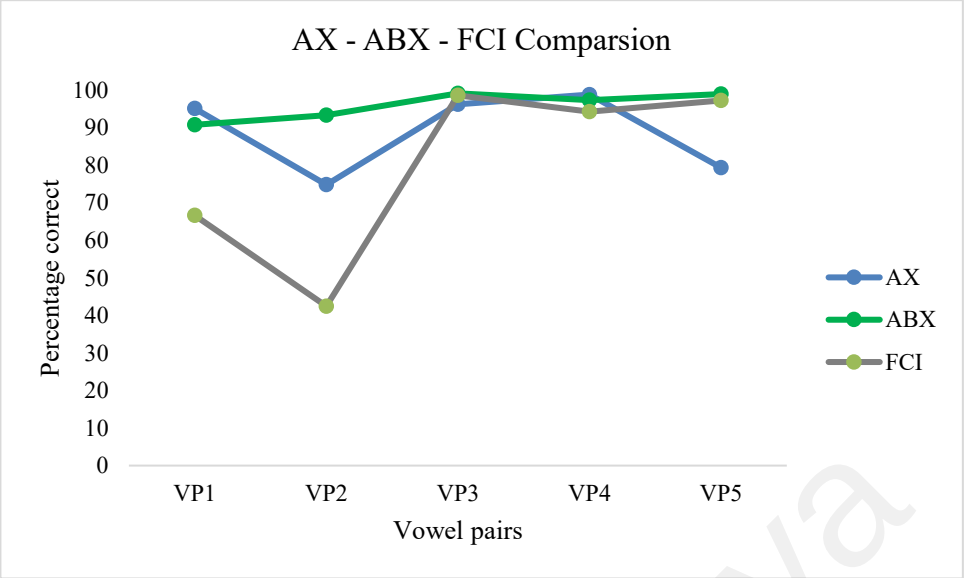

Figure 6.16: Comparison of all perception task

For VP1, the Mauchly's Test of Sphericity was significant, and thus, the findings of Greenhouse-Geisser were used with *epsilon* ε = 0.609. The findings indicated that the three tasks were significantly different $F(2, 60) = 1.370$, $p < 0.05$. A post hoc comparison using a Bonferroni adjustment showed that AX was not significantly different from ABX ($p = 0.436$) but significantly different from FCI ($p < 0.05$). ABX was also significantly different from FCI ($p < 0.05$). These findings suggest that for VP1, both AX and ABX were in the same subset homogenous group while FCI was not. It can be said that speakers distinguished VP1 in AX and ABX better than FCI.

For VP2, the Mauchly's Test of Sphericity was significant, and thus, the results of Greenhouse-Geisser are used with *epsilon* ε = 0.538. The findings indicated that the three tasks were significantly different $F(2, 60) = 35.645$, $p < 0.001$. A post hoc comparison using a Bonferroni adjustment showed that AX was significantly different from ABX ($p < 0.001$) and FCI ($p < 0.05$). ABX was also significantly different from FCI ($p < 0.005$). These findings suggest that for VP3, the three tasks were not in the same subset

homogenous group. It means that speakers perform significantly different for VP3 across the three tasks. Thus, by looking at the average score in Table 6.8, it can be said that speakers distinguished VP1 in AX better than ABX and FCI, while ABX was better than FCI.

Table 6.8: Average perception accuracy for the three tasks

	Vowel Pairs	AX	ABX	FCI	
VP1	[ɪ] – [iː]	95.16	90.81	66.67	84.21
VP2	[e] – [æ]	74.84	93.39	42.47	70.23
VP3	[ʊ] – [uː]	96.29	99.19	98.66	98.05
VP4	[ʌ] - [ɑː]	98.87	97.42	94.35	96.88
VP5	[ɒ] – [ɔː]	79.35	99.03	97.31	91.90

For VP3, the Mauchly's Test of Sphericity was significant ($p < 0.005$), and thus, the result of Huynh-Fledt was used with *epsilon* $\varepsilon = 0.814$. The findings indicated that the three tasks were not significantly different $F(2, 60) = 35.645$, $p = 0.077$. Since no difference was observed, no post hoc test was conducted for VP3. These findings suggested the speakers' perception score across the three tasks for VP3 was equivalent. Thus, no task type effect was found for this vowel pair.

For VP4, the Mauchly's Test of Sphericity was significant, and thus, the results of Greenhouse-Geisser were used with *epsilon* $\varepsilon = 0.672$. The findings indicated that the three tasks were significantly different $F(2, 60) = 4.486$, $p < 0.05$. A post hoc comparison using a Bonferroni adjustment showed that AX was not significantly different from ABX ($p = 0.321$) and FCI ($p = 0.067$). ABX was also not significantly different from FCI ($p = 0.236$). These findings suggested that speakers perform consistently higher across the three tasks for VP4.

For VP5, the Mauchly's Test of Sphericity was significant, so the results of Huynh-Feldt were used with *epsilon* ε = 0.946. The findings indicated that the three tasks were significantly different $F(2, 60) = 64.045$, $p < 0.001$. A post hoc comparison using a Bonferroni adjustment showed that AX was significantly different from ABX ($p < 0.001$) and FCI ($p < 0.001$). However, ABX was not significantly different from FCI ($p = 0.898$).

6.2 Production and Perception Relationship

To ascertain the relationship between the production and perception of English vowels by Acehnese and Indonesian speakers, two analyses were done. First the English vowels produced by the speakers were compared to the American speakers. The findings were then compared to the perception of English vowel pairs performed by the speakers. The comparison between the production and perception were reported descriptively. The second analysis involves the comparison of the production of every English vowel to equivalent Acehnese and Indonesian vowels. The findings were then compared to the perception of English vowel contrasts performed by speakers. The perceptual assimilation model (see 2.) was used to interpret the findings.

6.2.1 Comparison with the American Native Speaker of English

The English monophthong vowels, words used to elicit the vowel, average formant frequencies, and standard deviations for the F1 and F2 of each vowel produced by the American speaker can be observed in Table 6.9. Average formant frequencies were presented in both Hertz (Hz) and Bark scales. The Hz scale was used to describe the properties of each monophthong vowel relative to the others, while the Bark scale was used to plot the vowel into the vowel space area. The EDs (in Bark), or ED, are also presented in the far-right column. The average distance of all English vowels from the centroid [ɝ] is 1.79 Bark. This suggests that the English vowel space of the American speaker was more compact than the Acehnese male (2.7) and female (2.54). These results

may be due to the fact that the Acehnese speakers have more vowel system to accommodate as in Acehnese, Bahasa Indonesia, and English. This is in line with Vowel spaces were stretched in multilingual speakers in order to accommodate vowels from different languages and maintain their distinction (Best, 1994; Best & Strange, 1992).

Table 6.9: Formant values of English vowels by a male American speaker

Male	F1 (Hz)	F2 (Hz)	F1 (Bark)	F2 (Bark)	Euclidean
[iː]	310	2163	3	14	**3.52**
[ɪ]	458	1835	4	13	**1.80**
[e]	539	1800	5	12	**1.19**
[æ]	635	1688	6	12	**0.48**
[ʌ]	548	1455	5	11	**0.86**
[ɑː]	678	1387	6	11	**0.92**
[ɔː]	499	1004	5	9	**3.18**
[ʊ]	725	1761	7	12	**1.06**
[uː]	305	1307	3	10	**3.13**
[ɝ]	413	1461	4	11	**1.94**
Average					**1.79**

Figure 6.17 presents the scatter plot of English vowels produced by the American English speaker. The speaker was the same person who provided the audio used for the perception task of English vowel contrasts. It was apparent that most vowels are produced far apart from each other except the [ʌ] and [ɑː] vowel. Another interesting feature was the position of the [ʊ] vowel which was produced near the front-low vowel between the [e] and [æ]. The position of [e] and [e] was equally distant as [iː] and [ɪ]. The [ɔː] vowel was produced further apart from the [ɑː].

Based on the scatter plot of the American speaker's production in Figure 6.19, the relationship between Acehnese Indonesian speakers' production and perception of English vowel contrast could be determined. There were two ways of determining the relationship between perception and production by Acehnese Indonesian speakers. First was by classifying the distance among each English vowel pair produced by the American

speaker. Second was by comparing the production of English vowel contrast produced by American and Acehnese and Indonesian speakers.

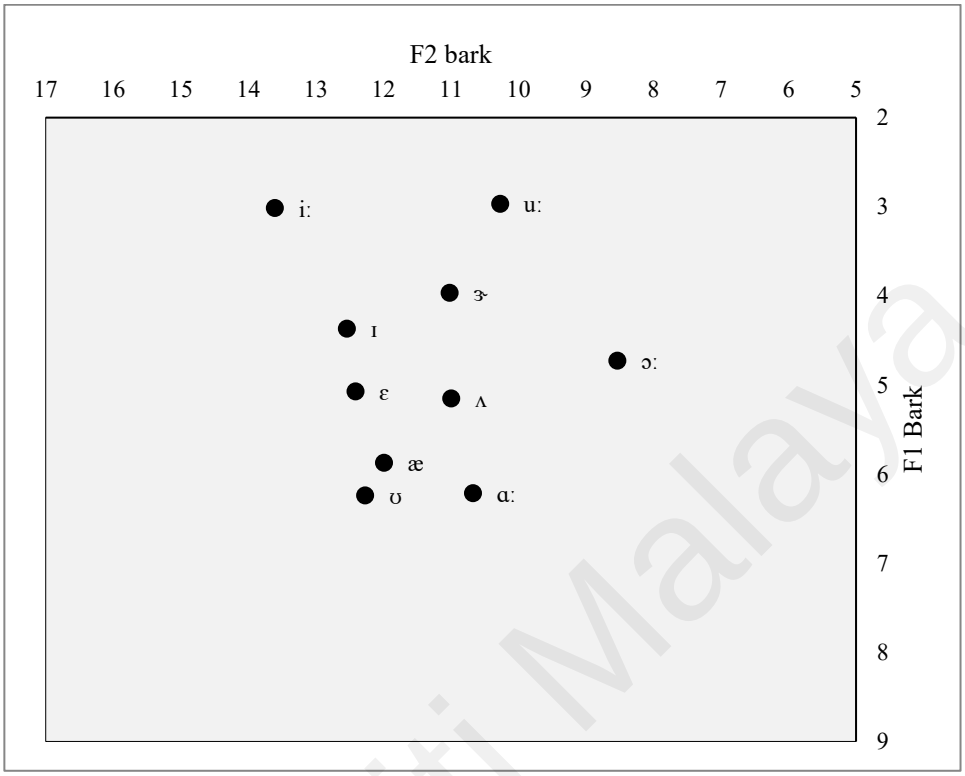

Figure 6.17: Production of English vowels by an American speaker

In the first method, the assumption was that the audio stimuli provided by American speakers plays a significant role in their ability to discriminate between the vowel contrasts. Their perception performance depends on the cues they heard from the audio stimuli. Thus, in this case, if the speakers relied on the cues heard from the American speaker's production not because of the assimilation of L3 English to the L1 and L2 sounds, the ease of discrimination of the vowel contrast would depend on the distant between the vowel pairs produced by the American in the scatter plot. Based on formant values and scatter plot produced by the American speaker, the possible difficulty the speakers face when distinguishing English vowel contrast could be categorized.

Based on formant values and scatter plot produced by the American speaker, we could categorize the possible difficulty the speakers face when distinguishing English vowel

contrast. Three vowel pairs [ʊ] - [uː] should be the easiest to distinguish since they are produced further apart from each other. Medium distinction could be observed for three vowel pairs, [ɪ] - [iː], and [ɑː] – [ɔː] while speakers might find it hardest to distinguish the [ʌ] - [ɑː] and [æ] - [ɛ] pairs.

In the second method, the speakers was assumed to rely on their production of English vowels where perception of English vowel contrast would be guided by their production of the said vowel contrasts. The audio stimuli provided by the American speaker may not have a greater impact on their ability to perceive the English vowel contrasts. Thus, the comparison of English vowels produced by the American English speaker and Ach-Ind male and female speakers was plotted in Figure 6.18 while the formant value of the Acehnese Indonesian speaker was presented in Table 6.10. It was obvious that Acehnese-Indonesian speakers occupied larger area of vowel space and which sat lower than the American speaker. This phenomenon was prominent among multilingual speakers. Vowel spaces were stretched in multilingual speakers in order to accommodate vowels from different languages and maintain their distinction (Best, 1994; Best & Strange, 1992). It was also apparent that Acehnese-Indonesian speakers in both male and female conflate some of vowel pairs while the American was not. Both [ɪ] - [iː] and [ʊ] – [uː] were produced closer together by the Acehnese-Indonesian speakers while the American speakers maintained distinction for both pairs. The American speaker on the other hand produced the [ɪ] - [iː] and [ʊ] – [uː] pair further apart from each other. Based on Acehnese Indonesian speakers' production of [ɪ] - [iː] and [ʊ] – [uː], it was expected that these two vowel pairs were difficult to distinguish.

Table 6.10: Formant values of English vowels by Acehnese-Indonesian speaker

Male	F1 (Hz)	F2 (Hz)	F1 (Bark)	F2 (Bark)	Euclidean
[iː]	371	2465	4	14	3.69
[ɪ]	362	2510	4	15	3.82

[e]	573	2168	5	14	2.16
[æ]	665	1971	6	13	1.52
[ʌ]	673	1522	6	11	0.40
[ɑː]	632	1194	6	10	1.84
[ɔː]	534	1072	5	9	2.68
[ʊ]	433	1094	4	9	2.95
[uː]	438	1056	4	9	3.12
[ɝ]	521	1485	5	11	1.00

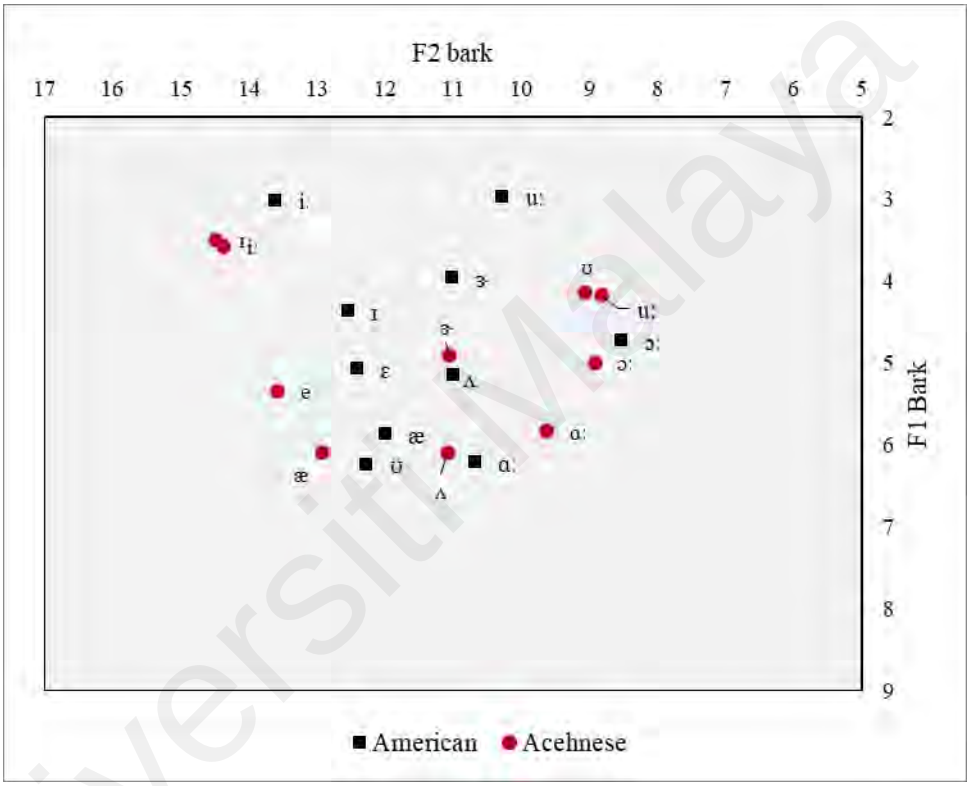

Figure 6.18: Comparison of the production position of American and Indonesian speaker

The Acehnese speakers produced the vowel pair [æ] – [e], [ɑː] – [ɔː], and equally distant from each other. The distance for [æ] - [ɛ] and [ʌ] – [ɑː] in Acehnese Indonesian speakers was equally like the American speaker. However, the American speaker produced the [ɑː] – [ɔː] pair much further than the Acehnese Indonesian speaker. On the other hand, the Acehnese speakers were producing the [ʌ] – [ɑː] pair much further apart compared to the American speaker. Both [ʌ] and [ɑː] are present in Acehnese so it was possible that the two vowels were assimilated into two existing Acehnese vowels. Further comparison of

English, Acehnese and Indonesian vowels could be observed in Section 5.4. This type of assimilation was also called a two-category assimilation in PAM term (Best & Tyler, 2007). Despite not being conflated, the American speaker obviously produced the pairs closely to one another. However, no statistical test could be conducted because the formant values of the American English vowels were provided only by one American speaker.

The comparison of distance of the five English vowel pairs between the Acehnese Indonesian and American speakers is presented in Table 6.11. English pairs under the *far* distance would be easy to discriminate in the perception task while those under the *medium* and *close* distance would be fairly difficult to distinguish.

Table 6.11: Prediction of English vowel contrast difficulty based on the American audio input

Vowel Pair	American Speaker	Ach-Indo Speaker
[ɪ] - [iː]	Medium	Close
[æ] - [ɛ]	Medium	Medium
[ʌ] – [ɑː]	Medium	Medium
[ɑː] – [ɔː]	Medium	Medium
[ʊ] – [uː]	Far	Close

In order to see if the perception of vowel contrast was bounded by the cue provided by the American speaker or by the production of the English vowel contrasts, this classification was compared against the perception score from AX, ABX and FCI task. The mean average of three perception tasks student took could be observed in Table 6.12.

Table 6.12: Mean average of three perception task, AX, ABX and FCI

	Vowel Pairs	**AX**	**ABX**	**FCI**
VP1	[ɪ] – [iː]	95.16	90.81	66.67
VP2	[æ] - [ɛ]	74.84	93.39	42.47
VP3	[ʊ] – [uː]	96.29	99.19	98.66

VP4	[ʌ] - [ɑː]	98.87	97.42	94.35
VP5	[ɑː] – [ɔː]	79.35	99.03	97.31

The first assumption was that the [ʊ] – [uː] pair should be the easiest to distinguish if the speakers rely on the cue from the American speaker but most difficult if they relied on their production of the vowel. This is because based on their production, the difference in [ʊ] – [uː] pair was not maintained and is classified as *single category* in PAM model which is difficult to discriminate in perception test. It can be seen from the Table 6.18 that the speakers scored above 90 for all the three tasks, AX, ABX and FCI indicating that this pair was easy to distinguish. For the [ʊ] – [uː] pair, speakers seemed to rely on the spectral cues found in the stimuli by the American speaker. Their difficulty in distinguishing the [ʊ] – [uː] pair did not affect their ability to discriminate this vowel. Another possible reason was that the final consonant of words used for this pair was also different, *foot* and *food*. The alveolar [d] was very apparent in the stimuli which might result in the speakers picking the cues beyond the spectral quality of the [ʊ] – [uː] pair.

The vowel pair [ɪ] – [iː] and [ɑː] – [ɔː] have a medium distance in the American speakers' production. In Acehnese-Indonesian speakers' production, the [ɪ] – [iː] was close while [ɑː] – [ɔː] was medium. If the speakers relied on spectral cue from American speakers' voice, both [ɪ] – [iː] and [ɑː] – [ɔː] should be medium to easy. If their production influences their perception, [ɪ] – [iː] pair should be more difficult than [ɑː] – [ɔː]. Speakers scored above 90 for [ɪ] – [iː] pair in AX and ABX tasks while for the FCI task, the score was 66.67. Similar pattern was found in the [ɑː] – [ɔː] pair, the score above 90 only recorded for ABX and FCI but not for AX task. Fr the AX task, the perception score of the [ɑː] – [ɔː] pair was 79.35. For [ɑː] – [ɔː] pair, discrimination of the vowel pair seemed to be influenced by the spectral cue provided by the American speaker. Even though the score for the [ɑː] – [ɔː] pair in AX task was below 90, at 79.35, the score was above

chance level suggested by Barrios et al. (2016). For the [ɪ] – [iː] pair, the result was mixed, the score 66.67 in FCI task was above chance level which also implied that this pair was difficult to distinguish in FCI task. However, the AX and ABX tasks, the speakers score higher which indicated they may rely in the cue provided by the American speaker. The speakers may be able to pick up the spectral cue and durational cue for this pair in the vowel perception task.

The distant for the [æ] - [ɛ] and [ʌ] – [ɑː] pair were of medium category in both American speaker Acehnese speaker. Since these two pairs were in the medium category, speakers' ability to discriminate between the two pairs in the perception task should be medium to easy. Looking the at the perception score, the perception score for [æ] – [ɛ] pair was above 90 only in ABX task while in AX and FCI task, the scores were 74.84 and 42.47 respectively. These figures indicated that the nature of the three tasks seemed to play a more significant role in the speakers' performance distinguishing the pair. However, the speakers were able to achieve higher scores in the ABX task indicating that the speakers were able to pick up the cue provided in the stimuli. The fact that Acehnese Indonesian speakers also distinguished the [æ] - [ɛ] pair in English production task may also contribute to their ability to score higher in ABX task. However, it was unclear why the score varied significantly across the three tasks. For the [ʌ] – [ɑː] pair, the score was above 90 in all three perception tasks. This uniform score across three tasks also resembled their production of the English vowel contrast. However, it was important to note that the words used for the [ɑː] vowel in *bard* had rhoticity due to American accent of the American speaker providing the stimuli. Thus, rhoticity might have contributed to a higher score in the perception task across the three tasks.

From the above analysis, there are some conclusions to be made. First, speakers' ability to perceive English vowel contrast was better than their ability to produce such vowels.

This finding was recorded for the [ɪ] – [iː] and [ʊ] – [uː] pairs. Speakers conflated both [ɪ] – [iː] and [ʊ] – [uː] pair in their production but managed to discriminate it well in perception task. They can perceivably distinguish the pairs but yet to materialize it into their production. Second, their ability to distinguish English vowel contrast was influenced by the spectral cues in the audio stimuli not by their ability to produce the vowel contrast. The spectral data for [ɪ] – [iː] and [ʊ] – [uː] pairs by the American speaker indicate that the two pairs were produced distant from each other. In fact, the [ʊ] – [uː] pair was produced almost twice the distance of [ɪ] – [iː]. In the perception task, the speakers distinguish the [ʊ] – [uː] pair above 90 across the three tasks. Third, nature of the three tasks play a significant role in their ability to discriminate English vowel pair regardless of the spectral cue in the stimuli. This finding was found in the [æ] - [ɛ] pair. The spectral distance of the pair produced by the American speaker was medium and the discriminate was expected to be medium to easy. However, the perception score was only higher in ABX task (above 90) but not in AX and FCI task. Similar variable patterns across three different tasks were also found in [ɪ] – [iː] and [ɑː] – [ɔː] pair.

6.2.2 Perceptual Assimilation Model

This section looks at how their production of English vowels relates to their production of Acehnese and Indonesian vowels. This comparison was done following the pred1iction made in PAM in which the assimilation of non-native vowels determined the ease of discrimination of non-native vowel contrasts. Instead of comparing the distance of English vowel pairs in the vowel space against the native speaker providing the stimuli, the assimilation of non-native English vowels into Acehnese and/or Indonesian produced by the Acehnese Indonesian speakers were compared. The assimilation is then classified PAM category, *single-category, two category* and *category goodness*. Each classification made different prediction on the ease of discrimination of English vowel contrasts in the

perception tasks. The assimilation of English pairs into Acehnese and Indonesian vowels were presented as follows.

Figure 6.19 indicates that all four vowels occupied the top left corner of the vowel space with all vowels conflated on top the others. It seems that the speakers produce the four vowels at the same F1 and F2 frequencies except few speakers who distinguished the four vowels. A repeated-measures ANOVA was conducted to confirm this assumption. For F1, the Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser was used with *epsilon* ε = 0.674. The findings indicated that the four vowels were not significantly different in term of F1, $F(3, 84) = 0.972$, $p = 0.385$. Since there was no difference was observed, no post hoc test was conducted for the F1 of these three vowels.

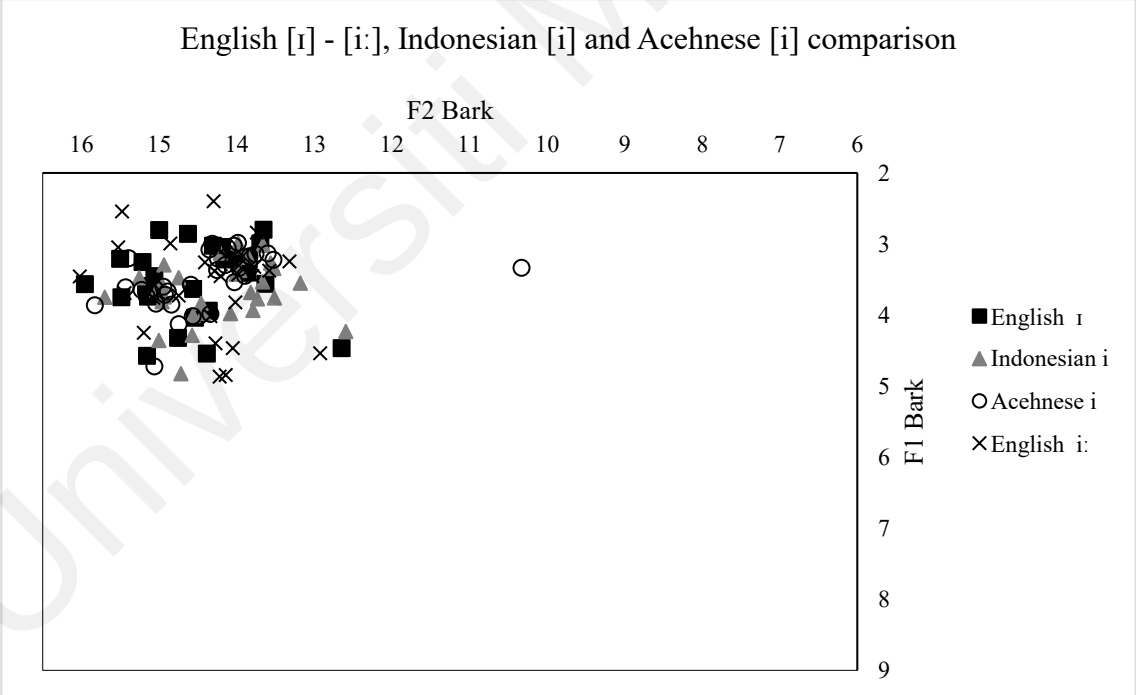

Figure 6.19: Comparison of English [ɪ] - [iː] with equivalent Acehnese and Indonesian vowels

For F2, the Mauchly's Test of Sphericity was significant, so the result of Hyunh-Feldt is used with *epsilon* ε = 0.870. The findings indicated that the four vowels were not significantly different in term of F2, $F(3, 84) = 1.370$, $p = 0.260$. Since there was no

difference was observed, no post hoc test was conducted for the F2 of these three vowels. Since both non-native contrasts were assimilated into both Acehnese and Indonesian [i], this vowel pair was classified as a *single category* assimilation in PAM term. The model predicts that the speakers should have difficulty in discriminating the English [ɪ] - [iː] in perception task.

Figure 6.20 shows that both English [ʊ] and [uː] were conflated to some Acehnese [u] but not to Indonesian [u]. However, both non-native vowels seemed to sit in between Acehnese and Indonesian [u], squeezing both native vowels apart to make ways for the nonnative vowels. A repeated-measures ANOVA was conducted to confirm this assumption. For F1, the Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser is used with *epsilon* $\varepsilon = 0.538$. The findings indicated that the four vowels were significantly different in term of F1, $F(3, 84) = 23.116$, $p < 0.001$. A post hoc comparison using a Bonferroni adjustment showed that F1 of English [uː] was significantly different from [ʊ] ($p < 0.001$), Acehnese [u] ($p < 0.05$) and Indonesian [u] ($p < 0.001$). English [ʊ] on the other hand was significantly different from English [uː] ($p < 0.001$) and Acehnese [u] ($p < 0.001$) but Indonesian [u] ($p = 1$).

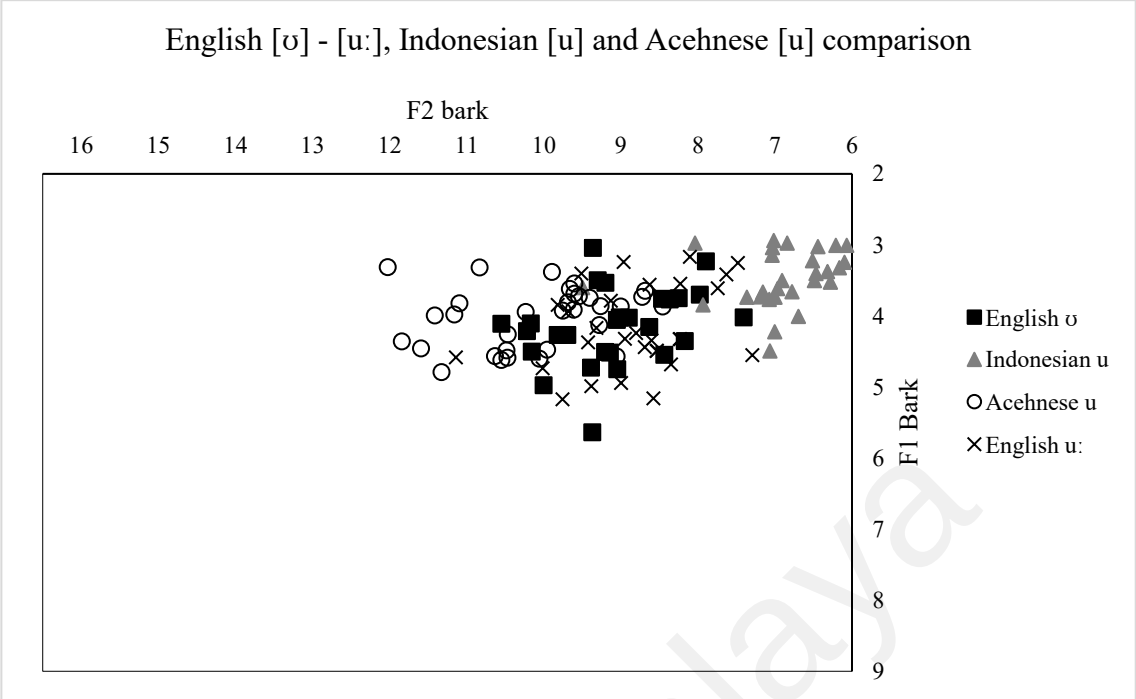

Figure 6.20: Comparison of English [ʊ] - [uː] with equivalent Acehnese and Indonesian vowels

For F2, the Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser was used with *epsilon* ε = 0.440. The findings indicated that the four vowels were significantly different in term of F2, $F(3, 84) = 14.69$, $p < 0.001$. A post hoc comparison using a Bonferroni adjustment showed that F1 of English [uː] was significantly different from Indonesian [u] ($p < 0.001$) but Acehnese [u] ($p = 1$) and English [ʊ] ($p = 0.509$). English [ʊ] was significantly different from Indonesian [u] ($p < 0.001$) but Acehnese [u] ($p = 0.34$) and English [ʊ] ($p = 0.509$).

Based on the statistical analysis, it indicated that only English [u] in which both F1 and F2 was significantly different from Indonesian [u]. English [ʊ] varied in terms of significance for F1 and F2 to either Acehnese or Indonesian [u]. Thus, it can be said that the quality of English [u] was similar to Acehnese [u] but different from Indonesian [u] but the quality of English [ʊ] similar to both Acehnese and Indonesian [u]. What interpreting using the PAM model is that the English [ʊ]-[uː] vowel contrasts can be

classified as a *single category* assimilation, in which both vowel contrasts were mapped into single native vowel. Discrimination of this vowel pairs in the perception task was predicted to be difficult by the PAM.

Figure 6.21 compares the English vowel contrasts [e] - [æ] to adjacent Acehnese [e]-[ə]-[ɛ] and Indonesian [e] - [ə]. It was apparent that the English [e] were mostly mapped into Indonesian [e] and Acehnese [ɛ]. Some English [æ] were produced closer to English [e], Indonesian [e] and Acehnese [ɛ] while some were further from existing Acehnese and Indonesian vowels. To confirm whether the English [e] - [æ] were conflated to adjacent Acehnese and Indonesian vowels a repeated measures ANNOVA was conducted. All seven vowels were compared against another to see which vowel was significantly different from another.

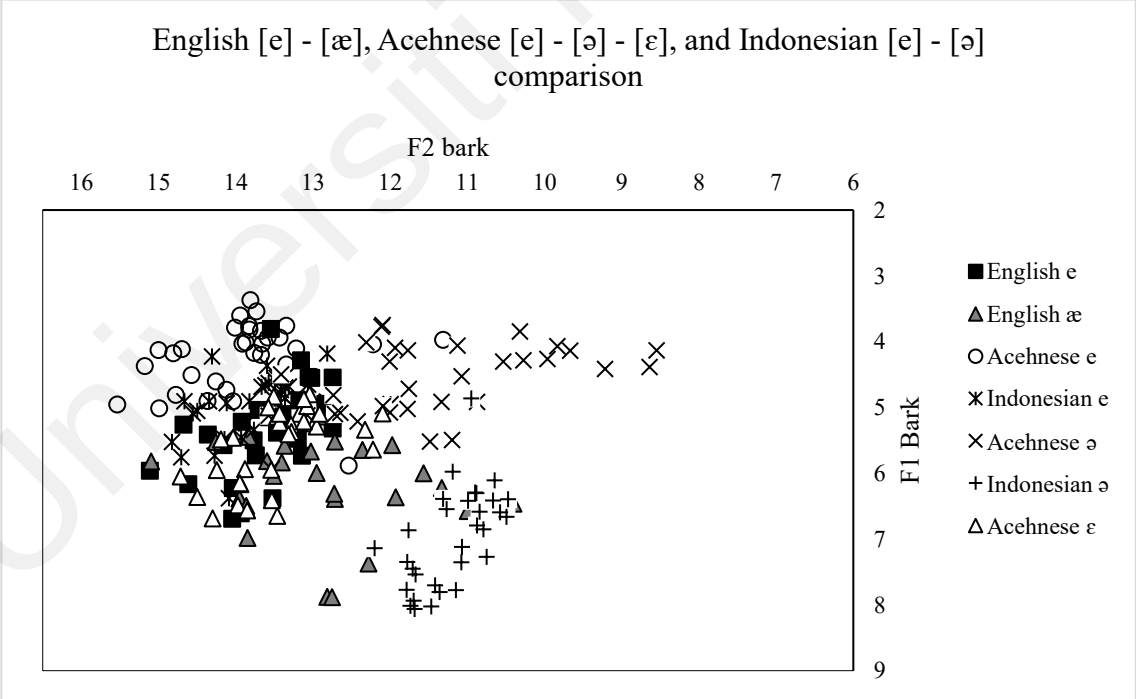

Figure 6.21: Comparison of English [e] - [æ] with equivalent Acehnese and Indonesian vowels

For F1, the Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser is used with *epsilon* $\varepsilon = 0.684$. The findings indicated that the seven vowels were

significantly different in term of F1, $F(3, 84) = 79.74$, $p < 0.001$. A post hoc comparison using a Bonferrroni adjustment showed that F1 of English [e] was significantly different from English [æ] ($p < 0.05$), Acehnese [e] ($p < 0.001$), Indonesian [e] ($p < 0.001$), Acehnese [ə] ($p < 0.001$) but Acehnese [ɛ] ($p < 0.12$) and Indonesian [ə] ($p = 1$). English [æ] was significantly different from English [e] ($p < 0.05$), Acehnese [e] ($p < 0.001$), Indonesian [e] ($p < 0.001$), Acehnese [ə] ($p = 0.001$) and Acehnese [ɛ] ($p < 0.05$) and Indonesian [ə] ($p < 0.001$).

For F2, the Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser is used with *epsilon* $\varepsilon = 0.440$. The findings indicated that the seven vowels were significantly different in term of F1, $F(3, 84) = 155$, $p < 0.001$. A post hoc comparison using a Bonferrroni adjustment showed that F1 of English [e] was significantly different from English [æ] ($p < 0.001$), Acehnese [e] ($p < 0.001$), Indonesian [e] ($p < 0.001$), Acehnese [ə] ($p < 0.001$) Acehnese [ɛ] ($p < 0.05$) but Indonesian [ə] ($p < 0.05$). English [æ] is significantly different from English [e] ($p < 0.05$), Acehnese [e] ($p < 0.001$), Indonesian [e] ($p < 0.001$), Acehnese [ə] ($p\ B = 0.001$) and Acehnese [ɛ] ($p < 0.05$) and Indonesian [ə] ($p < 0.001$).

Based on the statistical analysis, it indicated that the quality of English [ɛ] was similar to Indonesian Acehnese [ɛ] and Indonesian [ə] in term of height (F1) but not in term of the depth (F2). Thus, their similarity was partial and due to their difference in depth, the three vowels could well be different in qualities. The quality of English [æ] was different from any adjacent Acehnese and Indonesian vowels. Neither the F1 and F2 English [æ] was similar to any existing Acehnese or Indonesian vowels. of It means that the speakers were able to establish a distinguished quality for English [æ] different from native Acehnese and English vowels. In term of PAM, the [e] - [æ] pair can be classified as *category-goodness* assimilation in which one vowel was assimilated to native vowels (English [e]

partially assimilated to Acehnese [ɛ] and Indonesian [ə]) while the other was different from any existing native vowels. Discrimination for this vowel pairs was predicted to be moderately difficulty but easier than the *single* category.

Figure 6.22 compares the production of English [ɑː] - [ɔː] with Acehnese [a] - [o] - [ɔ] and Indonesian [a] - [o] - [ɔ]. The most noticeable pattern was that both English [ɑː] and [ɔː] are scattered across the Acehnese and Indonesian vowels. However, English [ɔː] was more spread out compared to [ɑː], scattering all over the vowel space but still withing the boundaries of Acehnese and Indonesian vowel inventories. Most of the English [ɑː] was produced within the area of Indonesian and Acehnese [ɔ]. It was assumed that these vowels were conflated to Acehnese [ɔ]. However, due to uneven pattern, it was difficult to determine if English [ɑː] was relatively similar to any Acehnese [ɔ].

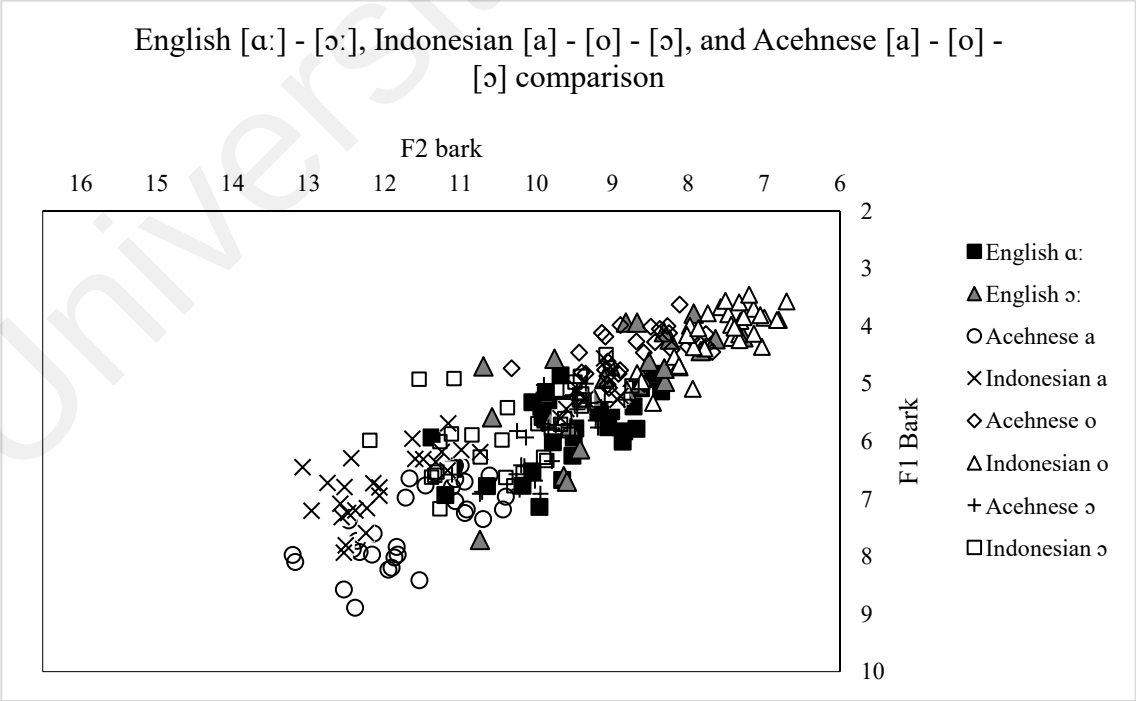

Figure 6.22: Comparison of English [ɑː] - [ɔː] with equivalent Acehnese and Indonesian vowels

Figure 6.23 on the other hand, compared English [ʌ] - [ɔ:] with Acehnese [a] - [o] - [ɔ] - [ʌ] and Indonesian [a] - [o] - [ɔ] to see if the English vowel pairs were conflated to any Acehnese or Indonesian vowels. The only difference in Figure X from Figure X was that in the current figure, an Acehnese [ʌ] was added to the equation and English [ɑ:] was replaced with English [ʌ]. What surprising in this figure was that English [ʌ] was nothing like Acehnese [ʌ]. English [ʌ] was presumably produced as instances of multiple different vowels within the Acehnese and Indonesian vowel inventories. A repeated measure ANOVA was run across 10 vowels, English [ɑ:] - [ʌ] - [ɔ:], Acehnese [a] - [o] - [ɔ] - [ʌ] and Indonesian [a] - [o] - [ɔ], and to see if any of English vowels were significantly different or similar to any comparable Acehnese and Indonesian vowels.

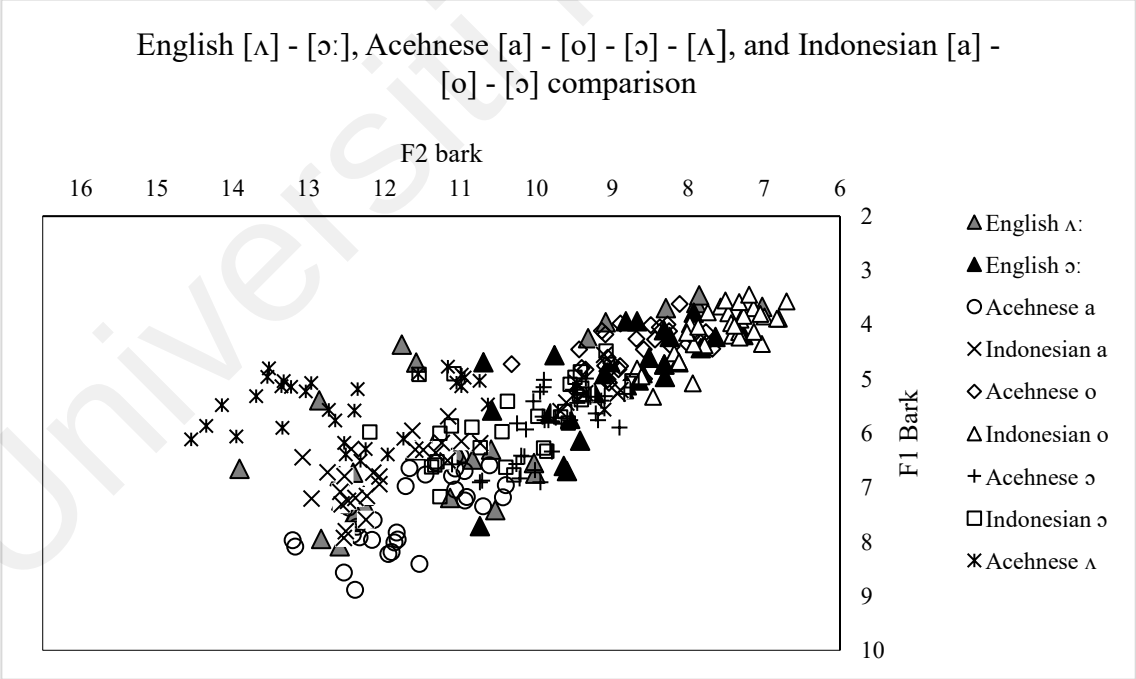

Figure 6.23: Comparison of English [ʌ] - [ɔ:] with equivalent Acehnese and Indonesian vowels

For F1, the Mauchly's Test of Sphericity was significant, so the result of Greenhouse-Geisser is used with *epsilon* ε = 0.256. The findings indicated that the ten vowels were

significantly different in term of F1, $F(9, 252) = 254.15$ $p < 0.001$. A post hoc comparison using a Bonferroni adjustment showed that F1 of English [ɑː] was significantly different from English [ʌ] - [ɔː], Acehnese [a] - [o], Indonesian [o] at $p < 0.001$ and Indonesian [a] at $p < 0.005$ but was not significantly different from Acehnese and Indonesian [ɔ] (p = 1.0) and Acehnese [ʌ] (p = 0.63). English [ɔː] was significantly different from English [ɑː] - [ɔː], Acehnese [a] - [ɔ], Indonesian [o] at $p < 0.001$ and Indonesian [ɔ] at $p < 0.005$ but is not significantly different from Acehnese [o] (p = 0.12) and Acehnese [ʌ] (p = 0.157). English [ʌ] was significantly different from English [ʌ] - [ɔː], Acehnese [a] - [o] - [ɔ] - [ʌ] and Indonesian [a] - [o] - [ɔ] at $p < 0.001$.

For F2, the Mauchly's Test of Sphericity was significant so, the result of Greenhouse-Geisser was used with *epsilon* ε = 0.283. The findings indicated that the ten vowels were significantly different in term of F1, $F(9, 252) = 56.13$, $p < 0.001$. A post hoc comparison using a Bonferrroni adjustment showed that F1 of English [ɑː] was significantly different from English [ʌ] - [ɔː], Acehnese [a] - [o] - [ʌ], Indonesian [a] - [o] at $p < 0.001$ but was not significantly different from Acehnese [ɔ] (p = 1) and Indonesian [ɔ] (p = 0.71). English [ɔː] was significantly different from English [ɑː] - [ʌ], Acehnese [a] - [ɔ] - [ʌ], Indonesian [a] - [o] at $p < 0.001$, Indonesian [ɔ] at $p < 0.05$ but not significantly different from Acehnese [o] (p = 1). English [ʌ] was significantly different from English [ɑː] - [ɔː], Acehnese and Indonesia [o] at $p < 0.001$, Acehnese [ɔ]- [ʌ] at $p < 0.05$ but was not significantly different from Acehnese [a] and Indonesian [a] - [ɔ] at p = 1.

Based on the statistical analysis, English [ʌ] - [ɑː] and [ɑː] - [ɔː] pair present some different pattern. The quality of English [ɔː] was similar to Acehnese [o] in term of height and depth. Thus, in both pair, English [ɑː] was assimilated both Acehnese and Indonesian [ɔ]. The English [ɑː], Acehnese and Indonesian [ɔ] vowels are produced within the same area of the vowel space. The quality of English [ɔː] was similar only to Acehnese [o] not

to Acehnese or Indonesian [ɔ]. English [ʌ] was significantly different from any Acehnese vowels. However, English [ʌ] was not an instance of different vowel from existing Acehnese and Indonesian vowels. The speakers produced this vowel variably across the vowel space in which same were being mapped into some Acehnese and Indonesian vowels. This is apparent in term of depth where the English [ʌ] were produced within the depth of Acehnese [a], Indonesian [a] and Indonesian [ɔ].

Perceptual assimilation model classifies vowel contrasts in which both vowels are assimilated into two different native vowels as *two-category* assimilation, which in this case is the English [ɑː] - [ɔː]. English [ɑː] was assimilated into Acehnese and Indonesian [ɔ] while [ɔː] was assimilated into Acehnese [o]. English Discrimination of this vowel contrasts is deemed the easiest among the PAM classification. Further elaboration on PAM could be observed in 2.5. English [ʌ] - [ɑː] could be classified as either *category-goodness* or *two-category* depending on how the degree of similarity of [ʌ] is interpreted. If having only F2 similar to Acehnese [a], Indonesian [a] and Indonesian [ɔ] was not counted, then the vowel contrast was classified as *category-goodness* assimilation. If this was the case, the vowel contrast should be perceived from good to very good depending on how strong the association of English [ʌ] to other Acehnese vowels. However, if the partial similarity of F2 to some Acehnese or Indonesian vowels is to be counted, the pair could be classified as a *two-category* scenario and should be discriminated easily in the perception task. However, the latter classification was more viable considering in PAM model, *category-goodness* was only applied if the English [ʌ] was not categorized into any of Acehnese or Indonesian vowels. Since some speakers produced English [ʌ] closer to multiple Acehnese or Indonesian vowels, it was assumed that the different speakers assimilated the vowel to different native vowels. The summary of the PAM classification for the five English vowel contrasts could be observed in Table 6.13.

Table 6.13 presents the PAM classification of the five English vowel pairs together with the average performance score of perception discrimination task for each pair. It was apparent that the prediction made by PAM did not seem to uniformly martialize in the perception task. *Single category* assimilation was predicted by PAM to be the most difficult pair to distinguish because the vowel contrasts are both realized as a single native vowel. This prediction did not hold true to both English vowel contrasts classified as single category assimilation, [ɪ] - [iː] and [ʊ] - [uː].

Table 6.13: Classification of English vowel pairs based on PAM in percentage

No	VP	PAM Classification	PAM Discrimination Prediction	Production	AX	ABX	FCI	Average
1	ɪ - iː	*Single-category*	Difficult	F1 – F2 conflated	95.16	90.81	66.67	84.21
2	e - æ	*Category-goodness*	Good to very good	No conflation	74.84	93.39	42.47	70.23
3	ʊ - uː	*Single-category*	Difficult	F1 conflated	96.29	99.19	98.66	98.05
4	ʌ - ɑː	*Two-category*	Easy	No conflation	98.87	97.42	94.35	96.88
5	ɑː - ɔː	*Two-category*	Easy	F1 – F2 conflated	79.35	99.03	97.31	91.90
			Average		88.90	95.97	79.89	

Despite being conflated to both Acehnese and Indonesian [u] in production test, the speakers were able to discriminate the [ʊ] - [uː] pairs at above 90 in the three discrimination tasks. This pair was also accumulated the highest average score at 98.05 compared to any other pairs tested. This pairs also yielded narrower standard deviation for AX (*M:* 96.20, *SD*: 7.06), ABX (*M:* 99.19, *SD*: 2.27), and FCI (*M:* 98.66, *SD*: 4.35). It means that speakers' performance was consistently higher across the three-discrimination tasks with score discrepancies among speakers lower. The fact that speakers were able to achieve excellent performance in perception task despite having

difficulty differentiating the pairs in the production task suggest that perception precede their production for this pairs. With an average of 98 percent across the three tasks, speakers achieved the highest score for this pair. With the average score of above 80 in all three tasks, this pair can be categorized as having good discrimination.

The production of both English [ɪ] and [iː] were also conflated into Acehnese and Indonesian [i], and yet speakers were still able to discriminate the pairs at above 90 in either AX or ABX task. However, poor discrimination performance was recorded for FCI task (*M*: 66.67, *SD:* 40.54). This poor average performance was mainly derived from the fact that the standard deviation for this pair was large in FCI task. This suggests that the average speakers' score for this pair more spread out. Detailed analysis of the data showed that five speakers obtained 0 score for this pair in the FCI task which lower the average score. However, it was also important to note that speakers average score was mostly lower than the other two tasks. With the average score of above 70 in the three tasks, this pair can be categorized as having fair discrimination.

Category-goodness assimilation pair ([e] – [æ]) was predicted to be discriminated from good to very good in perception tasks. However, the average score for AX and FCI tasks for this pair was the lowest among the pairs tested. The speakers only perform above 90 percent in ABX task. This was because speakers also score higher in ABX task for the other four pairs compared to AX and FCI task. The speakers especially score the lowest in the FCI task, which was below 50 percent. With the average score of above 70 in the three tasks, this pair can be categorized as having fair discrimination.

As for *two-category* assimilation, PAM predicted that the discrimination should the easiest among the assimilation types. Such prediction held true for the [ʌ] - [ɑː] but partially for the [ɑː] - [ɔː]. The speakers were able to discriminate the [ʌ] - [ɑː] pairs at above 90 percent across the three tasks. They were also able to achieve 90 percent average

score for the [ɑː] - [ɔː] in ABX and FCI tasks but failed to achieve similar score in AX task. The findings seem to be counter intuitive considering FCI got overall lower score in some other pairs, yet the speakers score higher in this task. Strong association of English [ɑː] to Acehnese and Indonesian [ɔ] did not seem to contribute to their performance in the AX tasks. However, with the average score of above 90 in the three tasks, this pair can be categorized as having good discrimination.

Based on the data from the Table 6.13, it is clear that the Perceptual Assimilation Model could not predict the ease of discrimination in all the five vowel contrasts. For example, the [ɪ] – [iː] pair is predicted to be difficult to discriminate yet the students were able to score above 90 percent except for FCI test. The [ʊ] – [uː] pair is also predicted to be difficult but the students scored 90 percent on all tests. Their assimilation prediction only materializes for the two-category assimilation and only for the [ʌ] - [ɑː] pairs not the [ʌ] - [ɑː] pairs. There were two possible contributions to these discrepancies, differences in the way the current study analysed for the assimilation pattern which was using the production data to determine assimilation patter, task type effect and possible other effect due to students made their choice based on cues other than the vowel. Thus, these results need to be treated with caution because some of the words used in the perception test does not qualify the minimal pairs and this becomes the limitation of the current study.

6.2.3 Second Language Linguistic Perception Model

Using the same comparison data from section 6.2.2, the findings were divided based on L2LP prediction. The review on L2LP model was presented in Section 2.5. The model made three scenarios regarding perception of non-native vowel contrasts based on the production of the said vowels. These prediction on perception accuracy was based on the assimilation of production of non-native sounds into native vowel system. The classification of L2LP prediction based on the current data into L2LP scenario was

presented in Table 6.14. English [ɪ] – [iː] and [ʊ] – [uː] were classified as *new* scenario (two non-native sound assimilated into single native sound) and they were supposed to be challenging to discriminate. Both English [ɪ] and [iː] were assimilated into Acehnese and Indonesian [i]. English [ʊ] and [uː] were both assimilated into Acehnese and Indonesian [u]. The rest of the pairs [ɛ] – [æ], [ʌ] - [ɑː], and [ɑː] - [ɔː] (two non-native sounds assimilated into more than two native sounds) were classified into *subset* scenario. All of these pairs were assimilated into multiple Acehnese and Indonesian vowels. Both new and subset scenario were predicted to be difficult to distinguish.

However, when looking at the English vowel contrast perception result, the prediction does not hold true. The [ɪ] – [iː] pair (*new*) were predicted to be challenging but the perception result in two of the three perception tasks was above 90. The challenging prediction was only occurred in FCI task, which was much more to do with the nature of the FCI task that makes speakers were difficult to distinguish the [ɪ] – [iː] pair. The [ʊ] – [uː] pair (*new*) was also predicted to be challenging to perceive but the findings do not confirm to the L2LP prediction. Speakers were able to distinguish the [ʊ] – [uː] pair in all three tasks at above 90. This suggests that the discrimination of this pair was easy despite the pair being assimilated into single native sound. The [e] – [æ] pair (*subset*) was predicted to be challenging and this prediction was true for AX and FCI task. However, the score of above 90 percent in ABX task indicated that the challenging situation resulted in AX and FCI task may be contributed by the difficulty of the perception task. For [e] – [æ] pair, the discrimination was challenging in AX and FCI but easy in ABX task. Another possible explanation was that speakers may change their perception across the tasks because this pair was in *subset* scenario where non-native sounds are assimilated into more than two native sounds.

Table 6.14: Classification of English vowel pairs based on L2LP in percentage

No	VP	L2LP Class.	L2LP Discrimination Prediction	Production	AX	ABX	FCI	Mean
1	[ɪ] – [iː]	*new*	challenging	F1 – F2 conflated	95.16	90.81	66.67	84.21
2	[ɛ] – [æ]	*subset*	challenging	No conflation	74.84	93.39	42.47	70.23
3	[ʊ] – [uː]	*new*	challenging	F1 conflated	96.29	99.19	98.66	98.05
4	[ʌ] – [ɑː]	*subset*	challenging	No conflation	98.87	97.42	94.35	96.88
5	[ɑː] – [ɔː]	*subset*	challenging	F1 – F2 conflated	79.35	99.03	97.31	91.90
			Average		88.90	95.97	79.89	

The [ʌ] – [ɑː] pair was also predicted to be challenging to discriminate in the perception task. This prediction was not observed in all three-perception tasks. Speakers were able to discriminate this pair at above 90 percent. The last pair predicted to be challenging is [ɑː] - [ɔː]. The discrimination of this pair is only challenging in AX task. In ABX and FCI tasks, the score was above 90 percent. This result suggested the data in the current study do not support the L2LP predictions. Speakers were all able to discriminate the five English pairs at above 90 percent in at least one perception task. It means that none of the L2LP's new and subset in this study was found to be challenging to be discriminated. Task design seemed to influence speakers' ability to discriminate the five English pairs.

6.3 Conclusion

This chapter presented the result for perception data and the relationship between the perception and production of vowels. The findings of the perception task show that the perception accuracy of each vowel pair was task dependent. The score for [ʊ] – [uː] and [ɑː] – [ɔː] were similar across the three tasks while the scores for the other pairs varied across the tasks. In general, speakers' discrimination score was mostly higher in the ABX, followed by AX and FCI tasks. However, the speakers were able to correctly discriminate

the five English vowel pairs above chance level in the three tasks except for [ɛ] – [æ] in the FCI task. Even though the speakers failed to discriminate [ɪ] – [iː] and [ʊ] – [uː] pairs in the production, they were able to distinguish the two pairs in the AX and FCI tasks. The difference was only observed in the ABX task where [ɪ] – [iː] was discriminated below chance level while [ʊ] – [uː] above chance level. The production of English [ɪ] and [iː] was not significantly different from Acehnese and Indonesian [i] while for English [ʊ] and [uː], there was a significant different from Indonesian [u].

CHAPTER 7: DISCUSSION

7.1 Introduction

This chapter discusses the findings presented in Chapter 4, Chapter 5 and Chapter 6 in relation to the research questions presented in Section 1.6. The current findings are discussed in relation to findings from previous studies and theories presented in Chapter 2. First, the quality of English monophthongs, quality contrast and length contrast are discussed in relation to existing studies on Indonesian context and neighbour countries in Indonesia. This part addresses the first (RQ1) and second research question (RQ2). Second, the comparison of quality of English monophthongs to existing Acehnese and Indonesian vowels are discussed by referring to PAM and L2LP models. This part addresses the third research question (RQ3). Finally, the perception results discussed descriptively followed by the discussion of production and perception relationship following PAM and L2LP models.

7.2 English monophthong quality

This section discusses the findings for the first research question: the quality of English monophthongs produced by Acehnese-Indonesian speakers based on their first and second formant measurements. The overall quality of the English monophthongs in the current study is, to some extent in terms of vowel space size and vowel spread, similar to the quality of English monophthongs produced by Acehnese speakers from Aceh Besar (Fata et al., 2017) and Javanese and Sundanese speakers (Perwitasari, 2019; Perwitasari et al., 2016). All speakers in these studies speak Indonesian as it is the national language. Speakers in the study conducted by Fata et al. (2017) and the current study are high school students from Aceh while speakers from Perwitasari (2019) study are university students. Speakers from Fata et al. (2017) study were from Aceh Besar and may speak a dialect of Acehnese from that region while in the current study the students are from West Aceh. The speakers from Perwitasari (2019) study have an intermediate level of proficiency in

English, which is on par with the speakers from the current study. No proficiency level was mentioned in Fata et al. (2017) study. The study conducted by Perwitasari (2019), Perwitasari et al. (2016) and Fata et al. (2017) also did not provide the data on ED so it is not possible to compare peripherality of the English vowels from the centroid. Thus, the comparison is made to other varieties in neighbouring countries.

The ED of English produced by Acehnese Indonesian speakers was larger than Singaporean English and Brunei English but slightly smaller than Malaysian English (Pillai, 2014). The vowel space area of the speakers in the current study was also bigger than the American English speakers providing audio stimuli for the perception task. The slightly larger ED of English vowels in these speakers seems to be influenced by the larger ED of their Acehnese and Indonesian vowels. This assumption is quite reasonable as it showed up as such when the English vowels they produced were compared to their production of Acehnese and Indonesian vowels. Most of these English vowels had a similar characteristic to both or either of the Acehnese and Indonesian vowels; this showed up in both F1 and F2 or either in F1 or F2. For example, English [iː] and [ɪ] were assimilated into Acehnese and Indonesian [i]. English [ɛ] was assimilated into Acehnese and Indonesian [ɛ]. English [ʊ] was assimilated into Acehnese [u]. This similarity is expected in L2 learners because they tend to filter non-native English vowels through their existing native vowels (Flege, 1995). Learners tend to weigh in the foreign sound and match the sound to any of the existing sounds in their vowel system. If they can discern the difference in the quality of non-native vowels to their existing native vowels, they can establish the phonetic category of that sound (Flege, 1995). On the other hand, if they fail to discern the difference between the quality of non-native sound and native sound, they will assimilate the non-native sound into the native sound. This is what happens to the English [iː] and [ɪ] in the current speakers. They failed to discern any difference in terms of quality for the two vowels which later had blocked the formation

of the new category. The assimilation of the non-native sounds to native sounds is expected in early L2 learners and this usually requires a challenging learning process to unlearn the already established assimilation (Escudero, 2005).

The Acehnese Indonesian speakers in the current study only conflated the English [ɪ] – [iː] and [ʊ] – [uː] while the Acehnese Indonesian speakers in the study carried out by Fata et al. (2017) conflated four English vowel pairs [ɪ] – [iː], [ʊ] – [uː], [ɛ] – [æ], and [ɒ] – [ɔː]. Despite having the same L1 and L2 with the speakers in Fata et al. (2017), the Acehnese speakers in the current study produced the [ɛ] – [æ] pair at a significant distance from one another. The English [æ] was produced lower than the English [ɛ] and this mostly happened with the male speakers. (see Figure 1.7, Section 4.2.3). However, the quality of Ach-Ind English [æ] is not necessarily of the same quality produced by the American and British speakers. This indication was spotted when the vowel was compared against Acehnese and Indonesian vowels produced by the same speakers (see Figure 2.11, Section 5.4.7). Some speakers approximated the English [æ] with the Acehnese and Indonesian [ɛ] and some with the Indonesian [a]. The approximation of the English [æ] to the Indonesian [a] mostly occurred with males in the current study. Significant difference in the English [ɛ] – [æ] pair for the current study was mostly due to the fact that some speakers produced the English [æ] similar to the Indonesian [a]. However, Fata et al's (2017) study did not compare the English vowel quality to Acehnese and Indonesian vowels produced by the same speakers, so it is unclear why the Acehnese Indonesian speakers in their study conflated this vowel. Javanese Indonesian and Sundanese Indonesian speakers also did not conflate the English [ɛ] – [æ], which suggests that categorical separation of this vowel pair might be common among Indonesian speakers (Perwitasari et al., 2016). Interestingly, Acehnese Indonesian speakers in the current study, Acehnese Indonesian speakers in Fata et al's (2017) study, and the Javanese Indonesian and Sundanese Indonesian speakers in Perwitasari et al's (2019) study

produced the English [ɛ] and [æ] in a relatively similar position. For all of these speakers, the English [ɛ] was positioned higher and more to the front than the English [æ]. This relative position of the English [ɛ] and [æ] in Indonesian speakers were similar to their production in Malaysian English in Pillai (2014) Pillai (2014). Brunei English speakers in another study (Sharbawi, 2006) produced the English [æ] further back but at a parallel horizontal position with the English [ɛ] instead of a parallel vertical position as with the speakers in the current study. In Singaporean English the [ɛ] and [æ] vowels were produced similar to the Acehnese Indonesian speakers in the current study. While categorical separation was maintained by both male and female Acehnese Indonesian speakers, the male Singapore English speakers did not maintain such a separation.

It is important to note that in Fata et al's (2017) study, the conflation of the English [ɒ] – [ɔː] was only found in male speakers. The female speakers on the other hand produced these two vowels further apart. However, the data for the English [ɒ] was not collected in the current study because the data collection was based on American English. Even though Fata et al's (2017) study was not explicit about the English variety used for data collection, it was assumed that they referred to British English because the English [ɒ] is commonly found in British English. Interestingly, the word used to target the British English [ɒ] in Fata et al's (2017) study and the American English [ɑː] in the current study was the same, using the word *pot*. In fact, the English [ɑː] in *pot* and [ɔː] in *port* were produced significantly differently by the Acehnese speakers in the current study. One reason for these discrepancies may be explained by different words used in this study and Fata et al's (2017) study. Fata et al's (2017) study used *bought* to capture the English [ɔː] while the current study used the word *port*. Rhoticity in the word *port* resulted in the vowel being pushed upward in the vowel space while the English [ɔː] in *bought* in Fata et al's (2017) study was produced lower in the vowel space.

Another interesting finding in the current study is the production of the English [ʌ] vowel in *bud*. The male and female speakers produced this vowel at a different location which suggests that they produced the vowel differently. The variety of the production was more apparent in the males than in females. Some male speakers produced the English [ʌ] vowel in *bud* like the Indonesian [u]. Instead of producing the vowel in *bud* with the [ʌ] vowel, some male speakers seemed to produce it as if it is an Indonesian word. As a result, the position of the vowel in *bud* among the male speakers in the vowel space was in between the English [ɝ] and [ɔː]. Therefore, the quality of the English [ʌ] produced by the male speakers does not reflect the actual quality of the vowel. In female speakers, on the other hand, the position of this vowel is more fronted and lower. This position of the English [ʌ] in the female speakers is similar to the male and female speakers in the study by Fata et al. (2017). A similar position for this vowel was also reported in Javanese Indonesian and Sundanese Indonesian speakers (Perwitasari et al., 2016). However, the Javanese Indonesian speakers conflated the English [ʌ] to the English [ɔː] while the Sundanese Indonesian speakers did not. The Acehnese Indonesian speakers in the current study and in Fata et al's (2017) study also did not conflate the English [ʌ] and [ɔː]. This suggests that the quality of the English [ʌ] produced by Acehnese Indonesian speakers is similar to Javanese but not to Sundanese. It would be interesting to see if other Indonesian speakers who speak different local languages produce the English [ʌ] similar to the Acehnese and Sundanese or similar to Javanese speakers.

The last feature worth discussing about the quality of English monophthongs produced by Acehnese Indonesian speakers is the rhoticity of two English vowels, [ɝ] and [ɔː]. Two of the words used to collect the production data involved the post-vocalic *r* in the words *bird* targeting English [ɝ] and *port* for the English [ɔː]. Both vowels in *bird* and *port* were produced at lower F2 values making the vowel less fronted. The r-coloured of vowels could be determined by lower F2 values (Hayward, 2000). This suggests that

Acehnese Indonesian speakers have rhoticity in their vowel production. Rhoticity has been reported as a feature of American English (Ladefoged & Johnson, 2014). Rhoticity was also a salient feature in Brunei English (Sharbawi & Deterding, 2010) and was also reported in a few speakers in Malaysian English (Gut & Pillai, 2014). Sharbawi and Deterding (2010) argued that rhoticity in Brunei English was mainly caused by Philippine English. As part of the American colony for 48 years until 1946, Philippine English has its origin from American English (Bautista & Gonzalez, 2006). Most Bruneian families employ helpers from the Philippines and these Filipino maids converse with the Bruneian children in an English influenced to an extent by American English (Sharbawi & Deterding, 2010). As a result, the Bruneian children grow up speaking English influenced to a certain degree by an American English pronunciation of words. For the speakers in the current study, it is a bit difficult to ascertain why rhoticity occurred. Indonesia was not colonized by the British as is the case with Malaysia and Singapore. It needs to be noted that Malaysian and Singapore English are not rhotic because these Englishes originated from British English. The only possible explanation is the popularity of American English media in Indonesia. Hollywood movies and American TV series are very popular in Indonesia. Before the disruption of internet media such as YouTube, major TV channels often aired Hollywood movies at night, promoting American English to millions of Indonesians. Popular internet video companies also originated from America such as YouTube, Facebook and Instagram. The speakers in the current study are Gen-Z and they grew up as digital natives consuming these popular internet video sharing platforms. Since the platforms originated from America, most of the early content creators were from America and spoke with an American accent. Another reason is that most English teachers in Aceh graduated from University of Syiah Kuala (Unsyiah) and Islamic University of Ar-Raniry (UIN Ar-Raniry), two of the largest and oldest universities in Aceh. Unsyiah has a very close relationship to the American Embassy and

used to have an American Corner to promote American English. However, now the American corner has been given to UIN Ar-Raniry. Most senior lecturers in both of these universities also graduated from American universities through the Fulbright scholarship, especially the post Tsunami Fulbright scholarship awarded to the Acehnese. The rhoticity in Acehnese Indonesian speakers also confirms the realization of the Kachruvian circle (Kachru, 1998). As the norm-dependent country, Indonesian tends to adopt the norms of English from the norm-providing country, that is, America.

7.3 Quality Contrast

This section discusses the findings from research question 2, the extent to which Acehnese-Indonesian speakers distinguish typical vowel pairs in terms of vowel quality and length contrast. It is known that English differs not only in quality but also in length. This section focuses only on the quality contrast while the discussion for durational contrast is presented in Section 7.3. The most interesting quality of English monophthongs by the Acehnese-Indonesian speakers in this study is the conflation of the English front vowels [iː] – [ɪ]. The overlapping of the English [iː] – [ɪ] was also reported by Acehnese-Indonesian speakers in the study by Fata et al. (2017) and in the English produced by Javanese and Sundanese speakers (Perwitasari, 2019; Perwitasari et al., 2016). It is important to note that neither study provided any statistical analysis on whether or not the two vowels are significantly different or similar. Their claims were made based on visual analysis of the two vowels in the vowel space. Fata et al. (2017) did mention the need for statistical analysis in further studies while Perwitasari (2019) was more concerned on comparing the English vowels against native speakers instead of vowel contrasts. However, the visual representation of [iː] – [ɪ] in both studies showed a similar resemblance to the current study in terms of conflation. Thus, it can be said that regardless of local languages, the Indonesian speakers do not contrast the [iː] – [ɪ] pair and produce them as the same vowels. The tendency of Indonesian speakers to conflate

the [iː] – [ɪ] pair seems to be contributed by the fact that Indonesian only has one high front vowel [i]. Like Acehnese, other local languages in Indonesia also have only one high front vowel since these languages belong to the same Austronesian family (Britannica, 2018).

The conflation of the high front vowel [iː] – [ɪ] is also an emergent feature of some Englishes in Asia. Lack of quality contrast for [iː] – [ɪ] was also reported in Malaysian English (Pillai et al., 2010), Singapore English (Deterding, 2003), ThaiE (Pillai & Salaemae, 2012), and Brunei English (Sharbawi, 2006). Thus, it can be argued that Indonesian English seems to conflate these two vowels just like other English varieties in ASEAN countries. This finding is interesting considering English is a foreign language in Indonesia unlike in Singapore, Malaysia and Brunei. Low (2010) classifies Singapore, Malaysia and Brunei as outer circle countries and Indonesia as an expanding circle country. Low (2010) also showed in her study that as norm dependent countries, expanding circle countries share features not only with the inner circle countries but also with expanding circle countries. In her study, the conflation of some English vowels such as [iː] – [ɪ] is emergent in either outer circle countries (Singapore and Philippines) or expanding circle countries (China & Indonesia). However, it is important to note that in her study, the Indonesian speakers live in Singapore and might have acquired the Singapore English features. In the case of the current study, the speakers have not lived in English-speaking countries or visited any neighbouring countries (Singapore and Malaysia). Thus, they have not had enough English input (either from the inner circle or expanding countries) other than from their friends and teachers (who are also Indonesian). The conflation is expected regardless of constant use of English as the language of communication in their school. When producing [ɪ] – [iː] pairs, which do not exist in their languages, the speakers assimilate them to existing native vowels, which will be discussed later in this chapter.

However, a recent study on Malaysian English, Pillai (2014) found that the Malaysian speakers were able to distinguish these pairs quite well. In this study, the English [ɪ] was produced a little bit higher and more in front than the English [iː]. This difference might be attributed to the speakers in this study who mostly speak English as their L1 while in the study by Pillai et al. (2010), it was Tamil, Chinese and Malay as their L1. The separation of [iː] – [ɪ] was also reported in the more recent study by Sharbawi (2012). This was despite the fact that the respondents in both the earlier study by Sharbawi (2006) and the later one Sharbawi (2012) were university students in Brunei. Different levels of proficiency might contribute to the development of such separations of [iː] – [ɪ]. Thus, further studies on Indonesian speakers with a higher level of proficiency will shed more light on whether or not Indonesians are able to distinguish these vowels.

Another interesting finding is the English [uː] and [ʊ]. In the current study, the English [uː] is the highest back vowel which is also the case in Fata et al. (2017) and Perwitasari et al. (2016). Thus, it can be said that despite different local languages, Indonesian speakers produce English [u] as the highest back vowel while [ʊ] is produced in a more fronted position. However, the relative position of [uː] and [ʊ] is slightly different in the current study compared to previous ones. In the current study, both vowels are parallel in terms of height but differ only in terms of depth. However, the difference between height and depth was not significant. In Perwitasari et al. (2016), both the Javanese and Sundanese speakers produced the [ʊ] vowel slightly lower than [uː]. Similar to the current study, the Javanese and Sundanese speakers also produced the [uː] and [ʊ] further apart from each other. This did not happen in the Fata et al's (2017) study where the English [uː] and [ʊ] are produced similarly.

The difference in the relative position of [uː] and [ʊ] by the current Acehnese speakers from previous Acehnese speakers in Fata et al. (2017) might be caused by different levels

of English proficiency. Even though both speakers in Fata et al's (2017) study and the current study are high school students, they differ in terms of daily English usage. The speakers from their study were from a conventional public high school while in the current study, the speakers are from an Islamic boarding school. In a conventional public high school, English is only taught during school hours while in the boarding school, in addition to learning English as a school subject, they regularly speak English outside the classroom. English is one of the modes of communication in addition to Arabic at the boarding school. Thus, the exposure teachers and peers get to English is higher in the current study. Jia et al. (2006) also found that English learners with longer exposure to English input perform better in production and perception.

In relation to emerging English varieties, the position of [uː] as the highest back vowel was also reported in SgE (Deterding, 2003), ThaiE (Pillai & Salaemae, 2012) and MalE (Pillai, 2014; Pillai et al., 2010). However, in BrunE, it is the vowel [ʊ] that is located as the furthest high back vowel while the position of [uː] is more in the front than [ʊ]. None of the English varieties mentioned above except BrunE indicated any conflation of the English [uː] and [ʊ] as similarly found in Acehnese Indonesian speakers in Fata et al. (2017). SgE, MalE, and ThaiE all distinguish the quality of [uː] and [ʊ] at different degrees of quality. Similar to these emerging varieties, Indonesian English produced by Acehnese-Indonesians in the current study also discriminate [uː] and [ʊ]. Discrimination of the two vowels were also salient in Indonesian English produced by Sundanese and Javanese Indonesians. Thus, it can be assumed that Indonesian speakers with a high level of proficiency could eventually distinguish [uː] and [ʊ] in their production. However, the relative position of [uː] and [ʊ] by Acehnese-Indonesian speakers in the current study is more similar to SgE and ThaiE than MalE and BrunE. In MalE, vowel [ʊ] is far lower than [uː] while in BrunE it is slightly lower. In both SgE and ThaiE, the vowel [uː] and [ʊ] are almost similar in terms of height but differ in terms of depth.

A previous study on Indonesian English by Fata et al. (2017) also reported that the English [e] – [æ] pairs were not contrasted. However, this is not the case in the current study. Even though the current speakers also speak Acehnese and Indonesian, they were able to distinguish the [e] – [æ] quite well (Perwitasari, 2019; Perwitasari et al., 2016). The current speakers were able to discriminate the pairs both in terms of depth and height and the difference was statistically significant. Sundanese and Javanese speakers were also able to distinguish the pairs quite well which is in agreement with the current study. The Javanese and Sundanese from a previous study (Perwitasari, 2019) and Acehnese speakers from the current study produced the [æ] a little bit lower than [e], but the relative position of the two vowels is different among the three groups of speakers. Acehnese and Javanese speakers tend to produce the [e] more in the front than [æ] while Sundanese do otherwise.

A study by Low (2016) on an Indonesian speaker living in Singapore also supports the current finding. The Indonesian speaker whose first language is Indonesian is able to distinguish the pairs in the production quite well. Thus, the current finding confirms previous reports that Indonesian speakers were able to distinguish the [e] – [æ] pairs. However, discrepancies with the findings from Fata et al. (2017) might be attributed to a different level of English proficiency. The speakers in the current study speak English as a language of communication just as the speaker in the study by Low (2016), while in Fata et al. (2017) the speakers only learn English during school hours, presumably passive English, which is apparent in most Indonesian public schools.

When compared to the regional varieties of English, the relative position of the [e] – [æ] pairs is on par with that produced in Philippine and Indian English (Low, 2016). Assuming that American English is more popular in Indonesia than British English, this resemblance is expected considering the Philippines was a former US colony (Kirkpatrick

& Sussex, 2012). Categorical separation was not observed in MalE (Pillai, 2014; Pillai et al., 2010), SgE (Deterding, 2003; Low, 2016) and BrunE (Sharbawi, 2006) despite the three countries sharing the same roots of the Malay language and having a close geographical proximity with Indonesia. For these particular pairs, Indonesian English tends to converge with inner circle countries compared to extending circles which further validate the claims made by Low (2016) that norm dependent countries might resemble either expanding or inner circle countries. Thus, it can be said that Indonesian English with different local languages separate the production of the English [e] – [æ] but differs in the relative position of the vowel against each other.

The last feature worth mentioning is the separation of [ʌ] – [ɑː] in the current study. This categorical distinction was also supported by Fata et al. (2017) especially in the male speakers. However, the production of the two vowels were much closer together in the Javanese and Sundanese speakers (Perwitasari, 2019; Perwitasari et al., 2016). One possible explanation for this difference is that Acehnese has the [ʌ] vowel (Pillai & Yusuf, 2012) while Javanese and Sundanese do not (Perwitasari, 2019). The Acehnese speakers might have utilised the Acehnese [ʌ] to discriminate the two English pairs. Further discussion on the use of the native vowel to filter non-native vowel contrasts will be elaborated. The separation of [ʌ] – [ɑː] was also not observed in the Indonesian speakers in the study by Low (2016). Since Low (2016) did not mention whether or not the Indonesian speakers in her study speak languages other than English and Indonesia, it cannot be ascertained why the production resembles Perwitasari (2019) instead of the current study. Thus, for Indonesian English the production of the English [ʌ] – [ɑː] might vary considerably depending on whether or not their local languages have one or both the English vowels in their vowel inventories.

The separation of [ʌ] – [ɑː] is not the common feature of English in Malaysia (Pillai, 2014; Pillai et al., 2010), Singapore (Deterding, 2003), Brunei (Sharbawi, 2006) and Thailand (Pillai & Salaemae, 2012). In contrast, Chinese, Indian and Philippine separate this pair to some extent (Low, 2016) which is similar to what is found in the current study. Thus, Indonesian English by Javanese and Sundanese resembles the earlier variety while Indonesian English Acehnese resemble the letter. However, it is important to note that the [ʌ] vowels in the current study were also realized as instances of multiple vowels by the current speakers which might contribute to the average value of the vowel being further apart from the [ɑː] vowel. Furthermore, British English which is the main reference for English in Singapore, Malaysia and Brunei does not distinguish the two vowels in terms of quality but mainly in terms of length(Deterding, 1997).

Nevertheless, the separation of the English [ʌ] and [ɑː] in the current speakers were mainly caused by the male speakers producing the target vowel differently. The word used to target the English [ʌ] was *bud*. The alphabet *u* in the word *bud* is pronounced as [u] in Indonesian. Since the speakers were asked to read the carrier sentence containing the target word on the computer screen, some speakers ended up producing the alphabet *u* as the [u] vowel instead of [ʌ]. The vowel [ʌ] exists in the Acehnese vowel system as in the word *cöt* [tʃʌt] meaning 'vertical, hill' but the Acehnese language do not have the written form. To further support this argument, the word *bud* was also rated as unfamiliar by most of the speakers in the word familiarity test. Most female speakers, despite being unfamiliar with the word, might have associated the pronunciation of bud to more common words such as *but* and *shut*.

7.4 Length Contrast

This section also discusses research question # 2 which focuses on the extent to which Acehnese-Indonesian speakers distinguish typical vowel pairs in terms of vowel quality

and length contrast. It is known that English differs not only in quality but also in length. Thus, the description of durational contrast among regional varieties of English is to be expected. The current study on the Acehnese-Indonesian speakers revealed that the speakers were able to maintain durational difference for [iː] – [ɪ], [ʌ] – [ɔː], [ʌ] – [ɝ] but not for [ʊ] – [uː] and [ʌ] – [ɑː]. The durational difference among these pairs was significant indicating that they discriminate the vowel quite well. This means that even though both Acehnese and Indonesian vowel systems do not have durational difference, the speakers were able to establish such a difference in their production.

The study on English vowel duration by the Javanese and Sundanese speakers reported that the speakers produced slightly reduced English monophthongs ([iː], [ɜː], [ɔː], [e], [æ], and [ʌ]) compared to native American speakers (Perwitasari et al., 2015). However, since no comparison was made among the English vowel contrasts, it was not known whether the Javanese and Sundanese speakers distinguished the long and short vowels in some English pairs. However, since the Javanese and Sundanese speakers produced some English pairs at a slightly reduced duration, there is a possibility that if compared against each other, some vowel pairs would have been significantly different in terms of duration. Also, there is no data on durational contrast presented in Fata et al's (2017) study despite the study being focused on English monophthongs produced by Acehnese high school students.

When the durational difference was compared against the quality difference, some interesting patterns emerged. Acehnese-Indonesian speakers do not discriminate [ɪ] - [iː] in terms of quality but discriminate the pair in terms of duration. The duration for the [iː] vowel was consistently longer than the [ɪ] vowel. It seems that the presence of two vowels in the written form of the word *beat* for the [iː] vowel may give the speakers the hint to lengthen this vowel. This also may explain that while [ɪ] - [iː] are produced closely to

each other, the speakers were still able to discriminate the pairs quite well in the perception task especially for AX and ABX tasks. Discussions about the findings on perception would be presented in Section 7.6. Interestingly, the speakers were able to discriminate the [ʌ] – [ɑː] pair not only in terms of duration but also in terms of quality. The presence of the post-vocalic *r* in the word *bard* for the [ɑː] vowel might have contributed to the vowel being slightly longer than the [ʌ] vowel. This also confirmed the assumption about the quality difference for the [ʌ] – [ɑː] pair considering the fact that the [ɑː] vowel in *port* has rhoticity. In perception, the speakers discriminated the pair above 90 in all three tasks which indicates that they were able to pick up the cues in the audio to discriminate the [ʌ] – [ɑː] pair. Speakers seemed to pick up the rhoticity cue in the word *bard* produced by the American speaker in the perception task to distinguish it from the [ʌ] vowel.

Interestingly, even though the Acehnese Indonesian speakers did not distinguish the [ʊ] – [uː] pair in terms of length and quality, they were able to discriminate this pair at perfect score in all three perceptions tasks, AX, ABX and FCI. One possible explanation for this disparity is that the speakers were able to observe the cue needed to differentiate between the two vowels in the perception task. The American speaker produced the [uː] vowel noticeably longer than the [ʊ] in the audio stimuli for perception tasks, which results in the speakers using this cue to differentiate the two sounds. In addition to the length, the final consonant in *foot* and *food* was also obvious in the audio stimuli used for the perception task. There is a possibility that speakers also relied on the final consonant as the cue to distinguish the word pair. It would be interesting to see how the speakers distinguish this pair by controlling initial and final consonants as in *bit – beat*. As for the production tasks, the speakers might have treated the *foot – food* pair as having the same vowel quality since the target vowels in both words used the same alphabetic vowel. This assumption is reasonable considering some male speakers produced the vowel in *bud* with

the [uː] phoneme instead of [ʌ], indicating that the alphabet used to represent the [ʌ] phoneme in *bud* played a significant role in the speakers' production. The speakers' ability to perceive vowel contrast in the perception task was better than in the production task for this pair. This supports the argument put forward by Flege (1995) that perception precedes production.

Lastly, the durational pattern does not all follow the durational pattern in American English especially for the [ʌ] – [ɑː] pairs. The vowel [ɑː] is actually a long vowel but the current speakers produce this vowel significantly shorter than the [ʌ]. In American English the [ɑː] vowel is longer than [ʌ] (Ladefoged, 2003) and for British English (Deterding, 2007), the difference in [ʌ] – [ɑː] is realized in terms of length instead of quality. In Malaysian English (Pillai et al., 2010) the [ɑː] vowel is also longer than the [ʌ] vowel. The data from the Javanese speakers (Perwitasari et al., 2015) showed that the Javanese speakers produced the [ɑː] vowel longer than [ʌ] while the American speakers produced [ʌ] longer than [ɑː]. The characteristics of [ʌ] – [ɑː] in the current study seems to resemble the native speakers in in the study by Perwitasari et al. (2015). However, the behaviour of the current speakers producing the [ɑː] vowel shorter than [ʌ] need further clarification. Since the current study only analysed the short and long vowel from one set of words, it cannot be ascertained whether this behaviour would also happen if other words are used for the [ʌ] – [ɑː] pairs.

7.5 English vowels quality to equivalent Acehnese and Indonesian vowels

Some models (SLM, PAM, and L2LP) in speech learning agree that when learning a new language, L2 learners tend to filter the target language sound through their existing sounds (Alispahic et al., 2017; Best & Tyler, 2007). As a result, the production of non-native sounds, at least in their initial stage of development, may resemble some existing native vowels. This study argued that the production of English vowels is to some extent related

to existing Acehnese and Indonesian vowels. In other words, the quality of some English vowels may be similar to the existing Acehnese and Indonesian vowels, especially the English vowels that are heard as having a similar quality to native vowels.

The findings showed that two similar and novel English vowels were produced with the quality indistinguishable from Acehnese and Indonesian vowels. This indistinguishable conflation was observed for the English [iː] and [ɪ]. When projected into vowel space the English [iː] and [ɪ] are indistinguishable from the Acehnese and Indonesian [i]. The conflation was so apparent that even statistical analysis did not provide any significant difference. Thus, in realizing the two vowels, Acehnese-Indonesian speakers relied on their existing [i] vowel in Acehnese and Indonesian and failed to realize any quality difference in the two English vowels. The English [iː] is similar to the Acehnese and Indonesian [i], but since Acehnese and Indonesian only have one high-front vowel, [i], they mapped the novel English [ɪ] together with the English [iː] to the Acehnese and Indonesian [i]. In relation to this, Escudero and Williams (2011) argued that in the early stage of non-native language learning, learners tend to approximate non-native vowel sounds to the nearest L1 sounds. However, in distinguishing the two vowels in the perception task, the current speakers seem to be able to pick up the durational cues resulting in the production of the English [iː] significantly longer than the English [ɪ]. Durational cues might also have contributed to their ability to discriminate between the two vowels in the AX and ABX tasks. Similar findings were reported in Yunus and Pillai (2020) on Malaysian multilingual speakers with L1 Malay, L2 English and L3 German. In this study, they found that no quality and length difference was observed in Malaysian [i], English [ɪ], and German [ɪ] in terms of F1 and F2. The study also found that the Malaysian speakers produced the Malaysian [i], the English [iː] and the German [iː] with the same quality. In terms of length, Malaysian [i] share the same length to German [iː] but shorter than than English [iː].

In addition to the total conflation, partial conflation was also observed for some English vowels. Partial conflation was especially observed for the English [uː], [ʊ], [ʌ], and [ɑː]. The partial conflation could be divided into two types, total conflation of only Acehnese vowels, and partial conflation of both Acehnese and Indonesian vowels. Total conflation to the Acehnese vowel occurred in the vowel [ʊ]. The quality of the English [ʊ] is indistinguishable from the Acehnese [u] but is separated from the Indonesian [u]. This seems to suggest that despite having the similar [u] vowel, the Acehnese and Indonesian [u] are significantly different. This finding further validates the argument made by Bohn and Flege (1992) and Flege (1987) that despite sharing the same phonetic symbol, the quality of two sounds from two different languages may differ considerably. Thus, for the English [ʊ] the association was stronger to the Acehnese [u] compared to the Indonesian [u]. It is assumed that when perceiving English words containing [ʊ] vowels, Acehnese-Indonesian speakers weighed the vowel as similarly as they would do in Acehnese compared to the Indonesian [u]. Further validation is required by gathering more data on more English words containing this vowel to be produced by Acehnese-Indonesian speakers.

Partial conflation to both Acehnese and Indonesia vowels was observed in the vowel [uː], [ɑː], and [ʌ]. Partial conflation occurs when only one aspect of vowel quality (height or depth) is similar to Acehnese and Indonesian vowels. The English [uː] shares similar properties with the Acehnese and Indonesian [u] but the difference is in terms of depth, it is similar to the Acehnese [u] while in terms of height, it is similar to the Indonesian [u]. The Acehnese and Indonesian [u] do not share the same properties in terms of height and depth. A similar scenario was also observed for the English [ʌ]. What was surprising is the English [ʌ] is significantly different from the Acehnese [ʌ] and in fact the speakers partially assimilate the vowel into the Acehnese and Indonesian [a]. One possible explanation for this is that the dialect of the current speakers which is the West Aceh

dialect, realizes the Acehnese [ʌ] differently from Acehnese dialect in Pase (Pillai & Salaemae, 2012). Different realizations of the Acehnese [ʌ] in the West Aceh dialect was first raised by Tanzir Masykar et al. (2021). In their studies, they found that the Acehnese speakers from West Aceh realized the Acehnese [ʌ] in the word *göt* as [ə], [ɛ], and [ɔ]. Such variations may result in the speakers having different properties of the Acehnese [ʌ]. What is not understood is why English [ʌ] is realized as the Acehnese and Indonesian [a] instead of any variation of the Acehnese [ʌ] from the West Aceh dialect. Thus, it is safe to assume that for the current speakers, the English [ʌ] has the same quality property with the Acehnese and Indonesian [a]. The speakers seem to partially borrow different properties from existing native vowels when realizing the English [uː] and [ʌ].

The English [ɑː] is novel to both the Acehnese and Indonesian vowel system. Since English [ɑː] is produced as a low back vowel in American and British English (Ladefoged, 2005; Ladefoged & Johnson, 2014), the novel vowel is to some extent realized as the mid back vowel of the Acehnese and Indonesian [ɔ]. They share the same quality in terms of depth but differ considerably in terms of height. This partial conflation indicates that the speakers may be aware that the English [ɑː] is novel to their vowel system but yet could not establish the vowel to de distinct from any existing native vowels. This assumption is made because the speakers discriminate the [ʌ] – [ɑː] [ɑː] – [ɔː] quite well in the perception task. The speakers might have picked up some cues in terms of the depth of the vowel, which is different from the depth of their existing vowel system. It would be interesting to see if this partial conflation will be fossilized or lost from existing native vowels in terms of height. Further studies on speakers with Advanced English proficiency may shed more light on this assumption.

In addition to total and partial conflation, three English vowels [e], [æ], [ɔː] were produced at significantly different quality from any existing Acehnese and Indonesian

vowels. Both the English [e] and [æ] were distinct from either the Acehnese and Indonesian [e] and [ə] while the English [ɔː] is different from either the Acehnese or Indonesian [ɔ]. This result is particularly interesting considering both the English [e] and [ɔː] have comparable vowels in Acehnese and Indonesia. The quality of the English [æ], which is novel to Acehnese and Indonesian vowels, is also different from any existing Acehnese and Indonesian mid vowels. It seems that the speakers have established the quality contrasts for the English [e] – [æ], not only between the two vowels but also between any existing native vowels. However, some degree of overlap is still observed in the English [e] and the Indonesian [ə]. This also explains why in terms of depth, the two vowels sit in the same location. The current speakers may be in the process of establishing a new category for some English vowels away from their existing vowels. This is particularly apparent when we look at the vowel space. The English [ɔː] was pushed up and to the back away from the Acehnese and Indonesian [ɔ] while the English [æ] is produced in between the Indonesian [e] and [ə].

7.6 Perception of Monophthong Contrasts

The findings from the perception data also present some interesting findings. The speakers' performance on the three perception tasks is not uniform and depends on the vowel pairs. Fluctuation in the performance across the three tasks indicate that some tasks might have been more difficult to solve than the others. The three tasks are of a different nature and require speakers to operate differently. Thus, inconsistencies would be expected.

Of all the three tasks, the FCI is the most difficult task and the average score for this task is among the lowest. Our assumption for FCI having the lowest score is that the task required the speakers to actively decide on the vowels they had heard instead of merely determining whether the vowels they heard were the same or different in the ABX and

FCI tasks. The FCI task was also administered at the end of the perception task. There is a possibility for the speakers to have found it to be a little harder to adjust to the FCI task which is very different in nature from the AX and ABX tasks.

Initially it was thought that speakers should find it easier to answer the AX task compared to the ABX task because the AX task is much simpler than the ABX task. In the AX task, the speakers only need to determine whether the two sounds they heard are similar or different while in ABX they have to decide if the last vowel they heard is similar to the first or second. However, the findings showed that the ABX task was easier than the AX task which suggests that the complexity of the ABX task did not hamper their performance. In fact, this design aided the speakers in making a better decision compared to the AX task. Actively deciding whether the X is similar to A or B seems to be more intuitive and allows speakers to retrieve better from their memory when evaluating the sounds heard. In the AX task, deciding whether the two sounds are similar or different might make them confused and reevaluate their decision. This was apparent when some speakers sometimes chose similar and then different sounds and vice versa. For example, in VP3, the speakers were able to identify the BB repetition but failed to do so for AA repetition even though they are both repetitions of the same vowels. This indicates that they got confused and had to re-evaluate their decisions throughout the task. In comparison, their score for AB and AX minimal pairs were higher than for repetitions (see Figure 6.5 in section 6.1.3).

Another interesting finding is that despite inconsistencies in the performance across the tasks, the speakers managed to score more than 90% in most pairs across the three tasks. If the threshold percentage was set lower at 70%, which is the minimum threshold set by Barrios et al. (2016) and Jia et al. (2006), then there were only two occasions when the speakers produced less than the threshold, [ɪ] – [iː] and [e] – [æ] for the FCI task. These

findings suggest that their performance in most of the pairs in all the three tasks is beyond near chance level. The speakers were very good at discriminating between the five English vowel pairs. If we look back at the data comparing the English, Acehnese and Indonesian vowels, it is clear that each vowel in the [ɪ] – [iː] pair is not significantly different from the Acehnese and Indonesia [i]. Thus, when the speakers had to decide which word they heard in the FCI task, they might have failed to realize the difference in quality. What is puzzling with the findings is that the speakers achieved a perfect score for AX and ABX tasks, above 90% but failed to do so in the FCI task. Even though in the production task, they failed to discriminate the two pairs [ɪ] – [iː], they managed to do so in the perception task. It is possible that in the AX and ABX tasks, the speakers were able to pick up durational differences in the words used for the [ɪ] – [iː] pair but failed to do so in the FCI task.

The findings for the [e] – [æ] pair also brought up the question as to why speakers still achieved a lower score for this pair despite the two vowels being significantly different in quality from one another and from any Acehnese and Indonesian mid vowels . In fact, the score for the [e] – [æ] pair in the FCI task is the lowest amongst the three tasks. The [e] – [æ] pair is also the only pair that has a score of less than 90% in two different tasks. Our assumption is that this was caused by the inability of the speakers to pick up the cues for durational difference. If we recall the findings of the durational contrast from the production data, it is obvious that the [e] – [æ] pair is the only pair in which speakers did not differentiate the durational contrasts. Since all other pairs differ significantly in terms of vowel length, this also led to our belief that durational contrast plays an important role in their ability to discriminate between each pair.

Looking at other factors, such as vocabulary size, does not seem to show any meaningful pattern in the speakers' performance. Classifying the speakers' vocabulary size into

small, medium, and large does not determine their achievement in the perception task. Speakers in the small, medium, and large vocabulary size groups all performed similarly across the three tasks for the five vowel pairs. These results are quite expected as the vocabulary size only determines their proficiency level in terms of vocabulary size but does not necessarily mean that their proficiency level in listening and speaking is the same.

7.7 Perception and Production

The last research question in the current study was aimed at looking at the extent to which the perception of English vowel contrasts is related to the production. Overall, the findings indicate that the speakers perceive the vowel contrast better than the production. On average, the speakers were able to discriminate the five English vowel contrasts at above chance level in the three perception tasks. For the production task, the speakers managed to discriminate the three vowel contrasts, [e] – [æ], [ʌ] – [ɑː], and [ɑː] – [ɔː] but failed to distinguish the [ɪ] – [iː] and [ʊ] – [uː] pairs.

Three pairs ([ʊ] – [uː], [ʌ] – [ɑː], and [ɑː] – [ɔː]) in the perception task had an average score of above 90% in all three tasks. This score resembles the score the native speakers got for [o] – [u], [ɑː] – [æ], [ɪ] – [iː] on the AX perception task in Barrios et al. (2016). Instead of relying on the difference between the target vowel pairs, the speakers may have relied on the production of adjacent consonants they heard in the target pairs. This also may explain why the average score for [e] – [æ] and [ɪ] – [iː] is lower than for the other three pairs. Both the [ʌ] – [ɑː] (bud – bard) and [ɑː] – [ɔː] (pot - port) pairs contain the alveolar trill /r/ in one of the words in the pairs. The American speaker produced the alveolar trill in the word “bard” and “port” as [bɑːrd] and [pɔːrt], which is not the case with British English. In British English, the words “bard” and “port” are produced as [bɑːd] and [pɔːt] respectively. American English often rhotacizes vowels followed by [r]

and is the norm in most parts of North America (Ladefoged & Johnson, 2014). The sounding of the alveolar trill [r] in one of the words in the pairs may have helped the speakers to discriminate between the two pairs at a near-native level. These two vowel pairs were also discriminated significantly in the production task. However, the difference in the [ʌ] – [ɑː] pairs appeared to be caused by variability in the production of [ʌ] by the speakers. The perfect score for the [ʊ] – [uː] may also have been caused by the production of the adjacent phoneme in the target words. In addition to lengthening the vowel [uː] in the word "food", the American speaker appears to stress the sound of the final consonant /d/. For the [ʊ] vowel, the American speaker shortened the vowel and the sound of the final consonant /t/ in 'foot' is apparent.

It is important to note the words used to collect the perception and production data for [e] - [æ], [ɑː] - [ɔː], and [ʌ] - [ɑː]. For the [e] - [æ] pairs, the words "beg' – 'bag" were used in the perception task while the words "bet' - 'bat" were used in the production task. Both pairs sit in the CVC context with the first consonant being a bilabial plosive and the last consonant being an alveolar plosive in production and a velar plosive in perception. For the [ɑː] - [ɔː] pair, the word "port" was used for the [ɔː] vowel in the perception task while the word "bought" was used in the production task. For the [ɑː] - [ɔː] pair, the word "bard" was used in the perception task and the word "pot" was used in the production task. Differences in the words used for the perception and production task might have contributed to the speakers discriminating better in the perception task than in the production task.

It was also assumed that the ease of discrimination of English vowel contrast in the perception task could be determined by comparing the production of English vowels by the Acehnese Indonesian speakers to the English vowels produced by the American speaker providing the audio stimuli for the perception task. Interestingly, speakers' ability

to discriminate between the five English vowel pairs was not related to their ability to produce the vowel. The speakers used the cues they heard from the audio stimuli produced by the American speaker to discriminate the English vowel contrasts.

The findings in the current study also did not fully support the PAM and L2LP models. There are three possible explanations as to why these discrepancies occurred. They are: (i) the nature of the three perception tasks used, (ii) the way assimilation was decided during the analysis, and (iii)the cues heard in the perception task. The three perception tasks used in the study which are the AX, ABX and FCI tasks had different levels of difficulty and the speakers performed variably in each task. For example, PAM has predicted that *single-category* assimilation would be difficult to discriminate in the perception task. The [ɪ] – [iː] pair in *bit – beat* is an example of the *single category*. In the L2LP model, this pair is classified as a new category where the discrimination is challenging. This classification was decided based on their production data where the speakers conflated the two vowels and assimilated these vowels into both the Acehnese and Indonesian [i]. In its original model, the classification was made by asking speakers to listen to non-native vowels and classify these vowels to their closest native vowels. However, in the perception task, the speakers were able to discriminate the English [ɪ] – [iː] pair in *bit – beat* above 90 in the AX and ABX tasks but below 70 for the FCI task. This indicates that the nature of the FCI task contributed to their poor performance in this task. Poor performance in the FCI task does not necessarily mean that as a *single category* pair, the [ɪ] – [iː] pair is difficult to discriminate because they successfully discriminated the same pair in the AX and ABX tasks at a perfect score. It needs to be noted that in the FCI task, the speakers were asked to listen to one word and decide on the list presented in the paper to which word the sound belonged to. In the AX and ABX tasks, they heard two and three sounds respectively which helped them in comparing the sounds they heard. Their good ability in discriminating in the AX and ABX tasks was

also driven by the cues they heard from the audio stimuli. The difference of *bit – beat* in the audio stimuli was apparent in terms of length but might not have been apparent in terms of quality for the Acehnese Indonesian speakers.

In another pair, the English [ʊ] – [uː] in *foot – food* was produced closely by the Acehnese Indonesian speakers and were partially assimilated into the Acehnese and Indonesian [u]. Because the two vowels were conflated, this pair is also classified as a *single category* in PAM and a *new* category in L2LP. Both models predict that the discrimination of the vowels would be difficult in PAM terms or challenging in L2LP terms. However, neither model was supported by the current findings in all three perception tasks. The speakers were able to discriminate this vowel pair at a perfect score above 90 points which would be considered easy. It was not their production of the vowel or assimilation of the non-native vowel into native vowels that assisted their ability to discriminate between the vowel pairs in the perception task but other factors such as the noticeable cue differences in the audio stimuli used in the perception task.

There were also instances of *two-category* assimilation (PAM) and *subset* category (L2LP) that partially supports the PAM model but does not support L2LP. The [ʌ] – [ɑː] and [ɑː] – [ɔː] pair are both classified as two-category PAM and subset category L2LP. As a *two-category* assimilation, PAM predicts that discrimination of the [ʌ] – [ɑː] and [ɑː] – [ɔː] pair in the perception task is easy. L2LP on the other hand predicts that as the *subset* category where the [ʌ] – [ɑː] and [ɑː] – [ɔː] pair is challenging to discriminate in the perception task. The findings indicate that both pairs were easy to discriminate in the three perception tasks except the [ɑː] – [ɔː] pair in the AX task thereby supporting the prediction made by the PAM model. Even though the average score for the [ɑː] – [ɔː] pair was below 90, its score was above 70 which suggests that the level of difficulty is in the medium range. Despite the result being inclined to support the PAM prediction for this

pair, other factors may also have played a role in facilitating speakers to discriminate between the vowels in the pair. It is because the words used for both pairs contain the post-vocalic r as in *bard* for [ɑː] and *port* for [ɔː]. Since the audio stimuli was from American English, the rhoticity of each vowel is apparent in both words. Speakers might have distinguished the [ʌ] – [ɑː] and [ɑː] – [ɔː] pair at a perfect score because they picked up the post-vocalic r in the target vowel. However, it is still unclear as to why the score for the AX task in the [ɑː] – [ɔː] pair is lower than the score in the other two tasks. One assumption is that the AX task is more difficult for speakers than the ABX task. However, the difficulty does not seem uniform based on the vowel pair. Some pairs were difficult to discriminate in the AX task while some others were difficult for the speakers in the FCI task.

7.8 Conclusion

This section has presented the discussions based on the findings . First, the quality of English vowels produced by the Acehnese Indonesian speakers were similar to those reported in previous studies by Indonesian speakers. This suggests that Indonesian English shares similar overall quality with minor differences found in the position of some vowels in the vowel space. Second, lack of quality contrasts was observed for [ɪ] - [iː] and [ʊ] – [uː] pairs which are similar to Malaysian English and Singaporean English. This suggests that Indonesian English has a similar pattern to other emerging Englishes in adjacent countries. Third, in terms of duration, the Acehnese speakers only maintain the durational difference for [iː] – [ɪ], [ʌ] – [ɔː], [ʌ] – [ɝ]. Fourth, the findings in the three perception tasks were not uniform across the vowel pairs. Some pairs were easy to discriminate in the AX task while some were easy in the FCI task. However, all three pairs were discriminated well in the ABX task which suggests that the ABX task was easier than the other two tasks. Lastly, their ability to discriminate the English vowel pairs in the perception task was assisted not by their assimilation of English vowels to

Acehnese and Indonesian vowels but by the cues they heard from the audio stimuli used in the perception task. Since the audio stimuli used American English, the post-vocalic r was apparent in the audio which may have been picked up by the speakers as the cues to distinguish the English pairs presented in the perception task.

CHAPTER 8: CONCLUSION

8.1 Conclusion

The current study has demonstrated the production and perception of English monophthongs by Acehnese-Indonesian speakers. Some interesting findings emerged in the current study, especially with regard to vowel contrasts. Their ability to discriminate vowel contrast was better in perception compared to the production task. Not all vowel pairs were discriminated equally in either the perception or production task. Some vowel pairs were discriminated easily in production, and some were discriminated easily in perception. Their ability to discriminate certain vowel pairs in the perception task does not necessarily reflect their ability to discriminate those pairs in the production task. For example, the [ʊ] – [uː] pairs were discriminated at a near-native level in the perception task even though the two vowels were conflated in the production task. The [e] – [æ] pairs on the other hand, were discriminated well in the production task, yet this achievement was not replicated in the perception task.

The study also found that some vowel conflation in the production task is influenced by the native vowel systems to a variable degree depending on the vowel pairs. Some English vowels were fully assimilated into existing Acehnese and Indonesian vowels while some were partially assimilated. For example, the current speakers conflated both the [ɪ] - [iː] and [ʊ] - [uː] pairs, but the two pairs were assimilated differently into Acehnese and Indonesian vowels. The quality of English [ɪ] and [iː] were indistinguishable from existing Acehnese and Indonesian [i]. The quality of the English [uː] was partially assimilated into the Acehnese and the English [ʊ] was also only assimilated into the Acehnese [u].

The overall quality of English vowels produced by the current speakers were similar to those reported in a previous study on high school students in Aceh (Fata et al., 2017).

However, differences were observed in regards to the number of vowels conflated. The current speakers conflated less vowels compared to the previous studies. In addition to conflating [ɪ] - [iː] and [ʊ] - [uː] as in the current study, the speakers in the previous study also conflated the [ɒ] – [ɔ] and [e] - [a] pairs. Different proficiency levels between the current speakers and the speakers in the previous study is assumed to have contributed to this disagreement. The conflation of the [ɪ] - [iː] and [ʊ] - [uː] pairs is also an emergent feature of English in Singapore, Malaysia, and Thailand. In relation to the English produced by Indonesian speakers from other regions, it seems that native local vowels from different languages in Indonesia contribute towards the production quality of some English vowels. The position of some English vowels produced by Acehnese-Indonesian speakers are slightly different from those produced by Javanese and Sundanese Indonesian speakers. Thus, Indonesian English may not be uniform across Indonesian speakers just as Indonesian dialects vary across the Indonesian archipelago.

8.2 Limitation

Despite some interesting findings, several limitations which were present in the current study prevent the researcher from making generalized conclusions on a larger population. First, some words used to collect the data for perception and production were different. Different adjacent consonants in the target words of some pairs may contribute to higher accuracy in the perception task compared to the production task. Differences in the words used were caused by the time gap between recording the audio for the perception tasks and data collection for the perception and production data. Audios for the perception tasks were provided by the American speaker but at the time of the data collection, the speaker went home to his home country which prevented re-recording the audios for the perception tasks. However, the environment in which the target vowel sits in the word is the same for the perception and production tasks.

Second, the number of speakers in the current study was less than 30 which is the minimum threshold required for some parametric statistical analysis. The number of speakers was initially 32 speakers, but three speakers were excluded from analysis because they missed some target English vowels in the production task. Therefore, the number of the speakers of this study was reduced to 29. However, the study proceeds with the t-test since simulation using wilcoxon-signed-rank, the comparable paired *t*-test for non-parametric test, yielded the same results. Furthermore, the values of Cohen-d test on some pairs was also small which is acceptable practice to use parametric-test.

Third, the nature of the perception task presentation may also contribute to speakers making pre-assumptions when doing the tasks. The randomized stimulus used in the perception tasks were made systematically across vowel pairs. It is possible that speakers who recognized the patterns in the first set of the pairs, made a pre-assumption judgement when listening to the subsequent pairs. However, despite this assumption, the results from the identification task and for certain vowel pairs do not present systematic guessing by the speakers.

Fourth, in addition to English, the speakers at Islamic boarding schools also learn Arabic and use the language alongside English. Thus, influences from the Arabic vowel system may be present in their speech. However, since most Acehnese can read the Holy Quran, it is highly likely that speakers who study in Islamic schools have learned the Arabic vowels since they were children. Being able to read the Holy Quran is a social demand in Aceh. The influence of Arabic vowels, because the speakers are able to read the Holy Quran, is inexorable among the Acehnese. Thus, it is recommended that future studies take into account any possible influences from other languages that participants may be fluent in.

Lastly, despite the fact that the Acehnese Indonesian speakers in the current study use English daily during their stay at the school, they lack access to authentic English input other than their teachers, seniors, and peers. This precondition may result in the speakers mispronouncing some words or guessing the pronunciation of the words based on the written forms. According to the findings, the pronunciation of some English words does not necessarily reflect the English spoken by other Acehnese Indonesian speakers in other settings. Some of the speakers pronounced *bud* by referring to the way Indonesian words are produced. Thus, using various words with different alphabets to represent the target vowel could help mitigate this limitation; this will aid in seeing if the speakers mispronounced the words using the written alphabet or are able to produce the correct pronunciation.

8.3 Suggestions

Future studies exploring English spoken by Acehnese-Indonesian speakers need to pay attention to the following aspects. First, methodologically, when collecting the data for perception and production of English vowels, one should decide on which English varieties to study, American, British or Australian English or even expanding English such as Malaysian and Singaporean English. Since the Indonesian curriculum does not explicitly mention which English varieties should be used when teaching English, it is possible that the options are up to the textbook author and classroom teacher. Words used to collect the data are very important especially if the study aims to capture different pronunciations that the study is focused on.

Second, when collecting perception data, future studies should use words that contain the same initial and final consonant to minimize the participants picking up cues other than the target vowels. This is especially important for the perception study since participants may be using cues other than the quality of the target vowels. However, this

may be difficult to achieve if American English is used as the reference point since some words have rhoticity. The other possible solution could be using more than two words for each pair to see if there are any differences in the outcome.

Third, studies on other groups of Acehnese who constantly speak English could also provide a more comprehensive picture on the quality of English produced by the Acehnese speakers. Future study could involve university students from the English department. Even though they do not speak English regularly, they learn English extensively throughout their study at the university. Sometimes they also speak English with their lecturers and friends. These university students are the future English teachers at schools and their English will eventually influence their future students . An alternative to English department students is English teachers at junior and high schools.

8.4 Implications

The current study could inform teachers and students alike when learning the pronunciation of English vowels. When designing a pronunciation course, the teacher could emphasize that some English vowel sounds are novel to Acehnese and Indonesian vowels and encourage the students to try to produce the novel vowels as intended and try not to conflate it with existing native vowels. This is especially important to the [ɪ] – [iː] and [ʊ] – [uː] pairs which have been reported to be conflated by Indonesian native speakers with various local languages like Acehnese, Javanese, and Sundanese.

This study also has theoretical implications in terms of emerging features of World Englishes. Studies on vowel quality of Asian English have mostly been focused on outer circle countries such as Singapore, Malaysia and Brunei and little data has been given to expanding circles such as Indonesia. This study has shown that Indonesian English does not seem to be uniform across the Indonesian archipelago. With hundreds of local languages spoken across the provinces in Indonesia, Indonesian English may owe its

features to these local languages. English vowels produced by Acehnese speakers appear to have a different quality to those produced by Javanese and Sundanese speakers. The English language features articulated by Indonesian speakers also depend on the speaker's level of proficiency. Students that use English regularly appear to be better at discriminating English vowel contrast which suggests that Indonesian English may not necessarily remain static to form a regional feature. It will constantly evolve as the speakers develop their level of English proficiency.

REFERENCES

Alispahic, S., Mulak, K. E., & Escudero, P. (2017). Acoustic properties predict perception of unfamiliar Dutch vowels by adult Australian English and Peruvian Spanish listeners. *Frontiers in psychology*, *8*, 52. https://doi.org/10.3389/fpsyg.2017.00052

Alwi, H., & Sugono, D. (2000). *Politik bahasa: Risalah seminar politik bahasa*. Pusat Bahasa, Departemen Pendidikan Nasional.

Amaro, J. C. (2012). L3 Phonology: An understudied domain. In J. C. Amaro, S. Flynn, & R. Jason (Eds.), *Third language acquisition in adulthood* (pp. 33-60). John Benjamins Publishing.

Antonina, L. (2017). Chapter five: Methodology: Choosing among paradigms and research designs. *Counterpoints*, *428*, 57-85. https://www.jstor.org/stable/45177753

Archibald, J. (2005). Second language phonology as redeployment of LI phonological knowledge. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *50*(1-4), 285-314.

Asyik, A. G. (1987). *A Contextual grammar of Acehnese sentences* [Doctoral dissertation, University of Michigan]. https://deepblue.lib.umich.edu/handle/2027.42/161507

Audacity Team. (2021). *Audacity(R): Free audio editor and recorder*. In (Version 3.0.0) [Computer software]. https://audacityteam.org

Aziz, E. A. (2003). Indonesian English: what's del tuh? *TEFLIN Journal*, *14*(1), 140-148. http://dx.doi.org/10.15639/teflinjournal.v14i1/140-148

Barrios, S. L., Jiang, N., & Idsardi, W. J. (2016). Similarity in L2 phonology: Evidence from L1 Spanish late-learners perception and lexical representation of English vowel contrasts. *Second Language Research*, *32*, 367-395. 10.1177/0267658316630784

Bautista, M. L. S., & Gonzalez, A. B. (2006). Southeast Asian Englishes. In B. B. Kachru, Y. Kachru, & C. L. Nelson (Eds.), *The Handbook of World Englishes*. Malden, MA : Blackwell Pub.

Be u , S. t. (2021). *Investigating spoken English : a practical guide to phonetics and phonology using Praat*. Palgrave Macmillan.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman; & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). MIT Press. https://doi.org/10.1007/978-94-015-8234-6_24

Best, C. T., & Strange, W. (1992). Effects of Phonological and Phonetic Factors on Cross-Language Perception of Approximants. *Journal of Phonetics*, *20*, 305-330.

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

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by a computer*. In (Version 6.3.10) [Computer software]. https://www.fon.hum.uva.nl/praat/

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in second language acquisition*, *14*(2), 131-158. https://doi.org/10.1017/S0272263100010792

Bolton, K., Botha, W., & Kirkpatrick, A. (2020). *The Handbook of Asian Englishes*. Newark : John Wiley & Sons, Incorporated.

Britannica, T. E. o. E. (2013). Malay language. In *Encyclopedia Britannica*.

Britannica, T. E. o. E. (2018). Austronesian languages. In *Encyclopedia Britannica*.

Britannica, T. E. o. E. (2020). German language. . In *Encyclopedia Britannica*.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (Vol. 1, pp. 4-64). Blackwell.

Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, *14*(2), 136-193. https://doi.org/10.1191/026765898669508401

Buschfeld, S. (2013). *English in Cyprus or Cyprus English: An empirical investigation of variety status*. John Benjamins Publishing.

Buss, S. (2014). *Get formant for vowels*. https://github.com/sbuss/get-formants-for-vowels/blob/master/get_formants_for_vowels.praat

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: complex or simple? Research case examples. *Journal of Research in Nursing*, *25*(8), 652-661. https://doi.org/10.1177/1744987120927206

Dardjowidjojo, S. (2009). *English phonetics & phonology for Indonesians*. Yayasan Pustaka Obor Indonesia.

Davydova, J. (2012). Englishes in the Outer and Expanding Circles: A comparative study. *World Englishes*, *31*(3), 366-385. https://doi.org/10.1111/j.1467-971X.2012.01763.x

De Angelis, G. (2007). *Third or additional language acquisition*. Multilingual Matters.

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

Deterding, D. (2003). An instrumental study of the monophthong vowels of Singapore English. *English World-Wide*, *1*, 1-16. https://doi.org/10.1075/eww.24.1.02det

Deterding, D. (2005). Emergent patterns in the vowels of Singapore English. *English World-Wide*, *26*, 179-197. https://doi.org/10.1075/eww.24.1.02det

Deterding, D. (2006). The pronunciation of English by speakers from China. *English World-Wide*, *27*(2), 175-198. https://doi.org/10.1075/eww.27.2.04det

Deterding, D. (2007). *Singapore English*. Edinburgh University Press.

Deterding, D., Jenkins, J., & Baker, W. (2013). *Misunderstandings in English as a lingua franca: An Analysis of ELF interactions in South-East Asia*. De Gruyter Mouton.

Deterding, D., & Kirkpatrick, A. (2006). Emerging South East Asian Englishes and intelligibility. *World Englishes*, *25*, 391-409. https://doi.org/10.1111/j.1467-971X.2006.00478.x

Durie, M. (1985). *A grammar of Acehnese on the basis of a dialect of north Aceh*. Foris Publications.

Errington, J. J. (1992). On the ideology of Indonesian language development: The state of a language of state. *Pragmatics*, *2*(3), 417-426.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands Graduate School of Linguistics.

Escudero, P., & Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *The Journal of the Acoustical Society of America*, *129*, EL1-L7. https://doi.org/10.1121/1.3525042

Fata, I. A., Fitrian, F., Mohammad, T., & Yusuf, Y. Q. (2017). Acoustic analysis on English oral vowels produced by Acehnese speakers from Aceh Besar by using Praat software. Proceedings of The 7th Annual International Conference (AIC) Syiah Kuala University and The 6th International Conference on Multidisciplinary Research (ICMR) in conjunction with the International Conference on Electrical Engineering and Informatics (ICELTICs) 2017, October 18-20, 2017, Banda Aceh.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65. https://doi.org/10.1016/S0095-4470(19)30537-6

Flege, J. E. (1991). The interlingual identification of Spanish and English vowels: orthographic evidence. *The Quarterly Journal of Experimental Psychology. Section A*, *43*(3), 701-731. https://doi.org/10.1080/14640749108400993

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). York Press. https://doi.org/10.1111/j.1600-0404.1995.tb01710.x

Flege, J. E., MacKay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, *106*, 2973-2987. https://doi.org/10.1121/1.428116

Flege, J. E., & MacKay, I. R. a. (2004). Perceiving vowels in a second language. *Studies in second language acquisition*, *26*, 1-34. https://doi.org/10.1017/S0272263104261010

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, *40*(4), 467-491.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /I/ and /l/ accurately. *Language and Speech*, *38*, 25-55. https://doi.org/10.1177/002383099503800102

Fokusaceh. (2020). *5 Pesantren Dayah Rekomendasi di Aceh Besar & Banda Aceh*. Retrieved July 27 from https://www.fokusaceh.com/2020/06/5-pesantren-dayah-rekomendasi-di-aceh.html

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America*, *97*(4), 2540-2551. https://doi.org/10.1121/1.411974

Gonçalves, B., Loureiro-Porto, L., Ramasco, J. J., & Sánchez, D. (2018). Mapping the Americanization of English in space and time. *PLoS ONE*, *13*(5), e0197741. https://doi.org/https://doi.org/10.1371/journal.pone.0197741

Gordon, R. (2005). *Ethnologue: Languages of the World*. SIL International.

Gregory, O. D. (1964). *Three Projects in English in Indonesia*. The Ford Foundation.

Grix, J. (2004). *Foundations of research*. Palgrave Macmillan Limited. http://ebookcentral.proquest.com/lib/latrobe/detail.action?docID=433562

Gut, U., & Pillai, S. (2014). Prosodic marking of information structure by Malaysian speakers of English. *Studies in second language acquisition*, *36*, 283-302. https://doi.org/10.1017/S0272263113000739

Hadisantosa, N. (2010). Insights from Indonesia. In R. Johnstone (Ed.), *Learning through English: Policies, challenges and prospects* (pp. 24-46). British Council.

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

Hayward, K. (2000). *Experimental Phonetics*. Longman.

Horne, E. C. (1961). *Beginning Javanese* (Vol. 3). Yale University Press.

Huda, N. (1999). *Language learning and teaching: Issues and trends*. Institute Keguruan dan Ilmu Pendidikan Malang

Hummel, K. M. (2013). *Introducing second language acquisition: Perspectives and practices*. John Wiley & Sons.

Ike, S. (2012). *Japanese English as a variety: Features and intelligibility of an emerging variety of English* [Doctoral Dissertation, University of Melbourne, School of Languages and Linguistics].

Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America*, *122*(5), 2842-2854. https://doi.org/10.1121/1.2783198

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, *126*(2), 866-877. https://doi.org/10.1121/1.3148196

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *The Journal of the Acoustical Society of America*, *97*(1), 553-562. https://doi.org/10.1121/1.412280

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Jenkins, J. (2006). Current perspectives on teaching world Englishes and English as a lingua franca. *TESOL quarterly*, *40*(1), 157-181. https://doi.org/10.2307/40264515

Jenkins, J. (2012). English as a Lingua Franca from the classroom to the classroom. *ELT journal*, *66*(4), 486-494. https://doi.org/10.1093/elt/ccs040

Jenkins, J. (2013). *English as a lingua franca in the international university: The politics of academic English language policy*. Routledge.

Jenkins, J. (2014). *Global Englishes: A resource book for students*. Routledge.

Jia, G., & Fuse, A. (2007). Acquisition of english grammatical morphology by native mandarin-speaking children and adolescents: age-related differences. *Journal of speech, language, and hearing research : JSLHR*, *50*, 1280-1299. https://doi.org/10.1044/1092-4388(2007/090)

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, *119*, 1118-1130. https://doi.org/10.1121/1.2151806

Kaboub, F. (2008). Positivist paradigm. In F. T. L. Leong, E. M. Altmaier, & B. D. Johnson (Eds.), *Encyclopaedia of counselling* (Vol. 2, pp. 343). https://go.gale.com/ps/i.do?p=GVRL&sw=w&u=latrobe&v=2.1&it=etoc&id=GALE%7C2FKN&sid=bookmark-GVRL

Kachru, B. B. (1985). *The alchemy of English : the spread, functions, and models of non-native Englishes*. Oxford : Pergamon.

Kachru, B. B. (1988). The sacred cows of English. *English Today*, *4*(4), 3-8. https://doi.org/10.1017/S0266078400000973

Kachru, B. B. (1992a). *The Other tongue: English across cultures* (2nd ed.). University of Illinois Press.

Kachru, B. B. (1992b). World Englishes: approaches, issues and resources. *Lang. Teach*, *25*(1), 1-14. https://doi.org/10.1017/S0261444800006583

Kachru, B. B. (1996). World Englishes: Agony and Ecstasy. *The Journal of aesthetic education*, *30*(2), 135-155. https://doi.org/10.2307/3333196

Kachru, B. B. (1998). English as an Asian language. *Links & letters*, 89-108. https://raco.cat/index.php/LinksLetters/article/view/22673

Kachru, B. B. (2005). *Asian Englishes Beyond the Canon*. Hong Kong China : Hong Kong University Press.

Kachru, Y., & Nelson, C. L. (2006). *World Englishes in Asian Contexts* (Vol. 1). Hong Kong University Press.

Kemendikbud. (2014). *Permendikbud No. 160*. Republik Indonesia

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.. ed.). Thomson Learning.

Khan, S. (2020). *10 Bahasa Daerah di Indonesia dengan Penutur Terbanyak, Ada Bahasamu?* Retrieved July 23 from https://www.idntimes.com/life/education/irma-yanti-2/penutur-bahasa-daerah-terbanyak-c1c2?page=all

Kirkpatrick, A. (2010). *English as a lingua franca in ASEAN: A multilingual model*. Hong Kong University Press.

Kirkpatrick, A. (2012). English in ASEAN: implications for regional multilingualism. *Journal of Multilingual and Multicultural Development*, *33*, 331-344. https://doi.org/10.1080/01434632.2012.661433

Kirkpatrick, A., & Sussex, R. (2012). *English as an international language in Asia: implications for language education* (Vol. 1). Springer Netherlands.

Kusnandar, V. B. (2022). *Indonesia Kini Punya 37 Provinsi, Cek Wilayah Sebarannya*. Retrieved 23 July from https://databoks.katadata.co.id/datapublish/2022/07/06/indonesia-kini-punya-37-provinsi-cek-wilayah-sebarannya

Ladefoged, P. (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Blackwell Pub.

Ladefoged, P. (2005). *Vowels and consonants*. Blackwell.

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics* (7th Edition ed.). Cengage Learning.

Lapoliwa, H. (1981). *A generative approach to the phonology of Bahasa Indonesia* [Doctoral Thesis, The Australian National University]. Canberra.

Larson-Hall, J. (2006). What does more time buy you? Another look at the effects of long-term residence on production accuracy of English vertical bar r vertical bar and vertical bar l vertical bar by Japanese speakers. *Language and Speech*, *49*, 521-548. https://doi.org/10.1177/00238309060490040401

Lauder, A. (2010). The status and function of English in Indonesia: A review of key factors. *Hubs-Asia*, *10*(1). https://doi.org/10.7454/mssh.v12i1.128

Lauder, A. F. (2020). English in Indonesia. In K. Bolton, W. Botha, & A. Kirkpatrick (Eds.), *The Handbook of Asian Englishes*. John Wiley & Sons, Incorporated.

Lee, D. D., & Low, E.-L. (2021). The sounds of Japanese English: Monophthong vowels and rhythmic patterning. *Asian Englishes*, *23*(1), 30-50. https://doi.org/10.1080/13488678.2020.1868815

Leimgruber, J. R. (2011). Singapore English. *Language and Linguistics Compass*, *5*(1), 47-62. https://doi.org/10.1111/j.1749-818X.2010.00262.x

Ling, L. E. (2010). English in Singapore and Malaysia. In A. Kirkpatrick (Ed.), *The Routledge handbook of world Englishes* (pp. 229).

Lipińska, D. (2015). Production of L3 Vowels: Is it Possible to Separate them from L1 and L2 Sounds? *Research in Language*, *13*(1), 77-92. https://doi.org/10.1515/rela-2015-0011

Llisteri, J., & Poch, D. (1987). Phonetic interference in bilingual's learning of a third language. Proceedings XIth ICPhS. The Eleventh International Congress of Phonetic Sciences,

Low, E. L. (2010). The acoustic reality of the Kachruvian circles: A rhythmic perspective. *World Englishes*, *29*, 394-405. https://doi.org/10.1111/j.1467-971X.2010.01662.x

Low, E. L. (2016). Phonological patterning for English as a lingua franca in Asia: Implications for norms & practice in multilingual Asia. *Journal of English as a Lingua Franca*, *5*, 309-332. https://doi.org/10.1515/jelf-2016-0022

Lowenberg, P. H. (1997). Testing English as a world language: Issues in assessing nonnative proficiency. In B. B. Kachru (Ed.), *The Other Tongue* (2nd Edition ed.). University of Illinois.

Mabruri, Z. K. (2021). Pengelolaan Bahasa Untuk Pembelajaran di Indonesia. *Naturalistic: Jurnal Kajian Penelitian Pendidikan dan Pembelajaran*, *6*(1), 1078-1085. https://doi.org/10.35568/naturalistic.v6i1.1609

Marsono, M. (2008). *Fonetik: Seri Bahasa*. Gadjah Mada University Press.

Masykar, T., Agusmaniza, R., Masykar, N. T., & Nurrahmi, F. (2021). An instrumental analysis of oral monophthongs in Aceh Barat dialect of Acehnese. *EduLite: Journal of English Education, Literature and Culture*, *6*(2), To be published. https://doi.org/10.30659/e.6.2.383-396

Masykar, T., Agusmaniza, R., Masykar, T., Shan, H., & Nurrahmi, F. (2021). Variation of Acehnese Monophthong/ʌ/in Western Acehnese Dialect. *Journal of English Language Teaching and Linguistics*, *6*(3), 695-714. https://dx.doi.org/10.21462/jeltl.v6i3.655

Masykar, T., Almulhim, A., & Nurrahmi, F. (2021). Diphthongized Monophthongs of Acehnese Oral Vowels in Samatiga Dialect. *Journal of Language and Literature*, *21*(2), Tobe published. https://doi.org/10.24071/joll.v21i2.3379

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, *30*, 229-258. https://doi.org/10.1006/jpho.2002.0174

McCulloch, L. (2005). *Aceh: then and now*. Citeseer.

McLellan, J. (2020). Brunei English. In K. Bolton, W. Botha, & A. Kirkpatrick (Eds.), *The Handbook of Asian Englishes*. Newark : John Wiley & Sons, Incorporated.

Meara, P. (1992). *EFL Vocabulary Tests*. Swansea University.

Mistar, J. (2005). *Teaching English as a foreign language (TEFL) in Indonesia*. Routledge.

Mohamad, N. R., & Deterding, D. (2016). The pronunciation of Brunei English. *World Englishes*, *35*, 519-528. https://doi.org/10.1111/weng.12223

Montolalu, L. R., & Suryadinata, L. (2007). National language and nation-building: The case of Bahasa Indonesia. In L. H. Guan & L. Suryadinata (Eds.), *Language nation and development in Southeast Asia* (pp. 39-50).

Mora, J. C. (2005). Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and nonwords by Spanish/Catalan bilingual learners of English. ISCA Workshop on Plasticity in Speech Perception (PSP2005),

Nailufar, N. N. (2022). *Masa Penjajahan Inggris di Indonesia*. Retrieved August 15 from https://www.kompas.com/skola/read/2020/02/11/140000669/masa-penjajahan-inggris-di-indonesia?page=all

Ozog, C. A. (1996). The unplanned use of English. In P. Martin;, C. Ozog;, & G. Poedjosoedarmo; (Eds.), *Language Use and Language Change in Brunei Darussalam* (pp. 156-172). OH: Ohio University Press.

Perwitasari, A. (2013). Slips of the ears: Study on vowel perception in Indonesian learners of English. *Jurnal Humaniora*, *25*(1), 103-110. https://doi.org/10.22146/jh.1816

Perwitasari, A. (2018). Perception of English vowels by Javanese and Sundanese speakers: A mouse-tracking study. *Wacana*, *19*(1), 219-234. https://doi.org/10.17510/wacana.v19i1.607

Perwitasari, A. (2019). *The acquisition of English vowels by Javanese and Sundanese native speakers* [Doctoral dissertation, Leiden University]. https://www.lotpublications.nl/the-acquisition-of-english-vowels-by-javanese-and-sundanese-native-speakers

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

Perwitasari, A., Klamer, M., Witteman, J., & Schiller, N. O. (2015). Vowel duration in English as a Second Language among Javanese learners. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, 2-5. https://doi.org/10.13140/RG.2.1.1752.5520

Pillai, S. (2014). The Monophthongs and Diphthongs of Malaysian English: An Instrumental Analysis. *English in Malaysia*, *35*, 56-85. https://doi.org/10.1177/003368820403500209

Pillai, S., Kaur, S., & Chau, M. H. (2021). The Ideological stance of multilingualism in education in Malaysia in the Press 2000-2020. *Austrian Journal of South-East Asian Studies*, *14*(2), 173-193. https://doi.org/10.14764/10.ASEAS-0058

Pillai, S., Mohd. Don, Z., Knowles, G., & Tang, J. (2010). Malaysian English: An instrumental analysis of vowel contrasts. *World Englishes*, *29*, 159-172. https://doi.org/10.1111/j.1467-971X.2010.01636.x

Pillai, S., & Salaemae, W. (2012). An instrumental analysis of English monophthongs produced by Thai EFL learners. *Pertanika Journal of Social Science and Humanities*, *20*, 1143-1160. http://www.pertanika.upm.edu.my/pjssh/browse/regular-issue?article=JSSH-0431-2011

Pillai, S., & Yusuf, Y. Q. (2012). An instrumental analysis of Acehnese oral vowels. *Language and Linguistics*, *13*, 1029-1050.

Platt, J., & Weber, H. (1980). *English in Singapore and Malaysia*. Oxford University Press.

Quinn, G. (2001). *What Is Indonesian? — The History of the Indonesian Language*. Retrieved 10 June from

Raditya, I. N. (2021). *Isi, Makna dan Sejarah Hari Sumpah Pemuda 28 Oktober 1928*. Retrieved October 24 from https://tirto.id/isi-makna-sejarah-hari-sumpah-pemuda-28-oktober-1928-eku2

Renandaya, W. A. (2000). Indonesia. In W. K. Ho. & R. Y. L. Wong. (Eds.), *Language Policies and Language Education: The Impact in East Asian Countries in the Next Decade* (pp. 113-137). Times Academic Press.

Riney, T. J., & Takagi, N. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*, *49*, 275-302. https://doi.org/10.1111/0023-8333.00089

Ringbom, H. (1987). *The Role of the First Language in Foreign Language Learning*. Multilingual Matters.

Roach, P. (2009). *English phonetics and phonology: a practical course* (Vol. 4th). Cambridge University Press.

Sadeghpour, M., & D'Angelo, J. (2022). World Englishes and 'Global Englishes': competing or complementary paradigms? *Asian Englishes*, *24*(2), 211-221. https://doi.org/10.1080/13488678.2022.2076368

Sadtono, E. (1997). ELT development in Indonesia: A smorgasbord. In E. Sadtono (Ed.), *The development of TEFL in Indonesia* (pp. 1-19). Institute Keguruan dan Ilmu Pendidikan Malang.

Sarumpaet, J. P. (1963). *Indonesian secondary schools*. Secondary Teacher.

Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. *Language*, *79*, 233-281. https://www.jstor.org/stable/4489419

Schneider, E. W. (2007). *Postcolonial English: varieties around the world*. Cambridge University Press.

Schneider, E. W. (2014). New reflections on the evolutionary dynamics of world Englishes. *World Englishes*, *33*, 9-32. https://doi.org/10.1111/weng.12069

Seidlhofer, B. (2004). Research Perspectives on Teaching English as a Lingua Franca. *Annual review of applied linguistics*, *24*, 209-239. http://ez.library.latrobe.edu.au/login?url=https://www.proquest.com/scholarly-journals/research-perspectives-on-teaching-english-as/docview/85671035/se-2?accountid=12001

Seidlhofer, B. (2013). *Understanding English as a lingua franca-Oxford applied linguistics*. Oxford University Press.

Sharbawi, S. (2006). The vowels of Brunei English An acoustic investigation. *English World-Wide*, *3*, 247-264. https://doi.org/10.1111/j.1467-971X.2011.01725.x

Sharbawi, S. (2012). Revisiting the vowels of Brunei English. *World Englishes*, *31*, 177-197. https://doi.org/10.1111/j.1467-971X.2011.01725.x

Sharbawi, S. (2021). English use in Brunei: Investigating the Gen Z. *Asian Englishes*, 1-25. https://doi.org/10.1080/13488678.2021.1961376

Sharbawi, S., & Deterding, D. (2010). Rhoticity in Brunei English. *English World-Wide*, *31*, 121-137. https://doi.org/10.1075/eww.31.2.01sha

Sobarna, C. (2007). Bahasa Sunda sudah di ambang pintu kematiankah? [Is the Sunda language at the brink of extinction]. *Makara Human Behavior Studies in Asia*, *11*(1), 13-17. https://doi.org/10.7454/mssh.v11i1.39

Soderberg, C., & Olson, K. (2008). Indonesian. *Journal of the International Phonetic Association*, *38*(2), 209-213. https://doi.org/10.1017/S0025100308003320

Subramaniam, P. (2008). *An acoustic analysis of Malaysian English vowels* [Master Thesis, Jabatan Bahasa Inggeris, Fakulti Bahasa dan Linguistik, Universiti Malaya]. https://doi.org/10.5539/ass.v10n14p254

Sundkvist, P., & Gao, M. (2016). Rhoticity in Yunnan English. *World Englishes*, *35*, 42-59. https://doi.org/10.1111/weng.12172

Sundkvist, P., & Nguyen, X. N. C. M. (2020). English in Vietnam. *The handbook of Asian Englishes*, 683-703.

Tan, E. B. (2007). Language policy and discourse in Singapore. In H. G. Lee & L. Suryadinata (Eds.), *Language, Nation and Development in Southeast Asia* (pp. 74-117). Institute of Southeast Asian Studies.

Thirusanku, J., & Yunus, M. M. (2014). Status of English in Malaysia. *Asian Social Science*, *10*(14), 254.

Thomas, R. M. (1968). Indonesia: The English-language curriculum. In R. M. S. Thomas, L.B. & D. L. Brubaker (Eds.), *Strategies for curriculum change: Cases from* (Vol. 13, pp. 279-322). International Textbook Company.

Tilaar, H. A. R. (1995). *Lima Puluh Tahun Pembangunan Pendidikan Nasional 1945–1995*. Grasindo.

Trudgill, P., & Hannah, J. (2013). *International English Fifth Edition* (5th ed.. ed.). Hoboken : Taylor and Francis.

Widagsa, R. (2015). Acoustic analysis of English vowel formant frequencies produced by Indonesian native speakers. The 2nd International Language and Language Teaching Conference, Yogyakarta.

Widagsa, R., & Putro, A. A. Y. (2017). Acoustic measurement on vowel production of English as a Second Language by Indonesian EFL Learners. *English Review: Journal of English Education*, *6*(1), 71-80. https://doi.org/10.25134/erjee.v6i1.772

Widiati, U., Rohmah, Z., & Furaidah. (2017). *Buku Guru Bahasa Inggris*. Pusat Kurikulum dan Perbukuan, Balitbang, Kemendikbud.

Wildan. (2002). *Tata Bahasa Aceh*. Global Educational Consultant Institute.

Wrembel, M. (2010). L2-accented speech in L3 production. *International Journal of Multilingualism*, *7*, 75-90. https://doi.org/10.1080/14790710902972263

Wrembel, M. (2012). Foreign accentedness in third language acquisition: The case of L3 English. In J. C. Amaro, S. Flynn, & R. Jason (Eds.), *Third language acquisition in adulthood* (pp. 281-309). John Benjamins Publishing.

Wrembel, M. (2013). Foreign accent ratings in third language acquisition: The case of L3 French. In *Teaching and researching English accents in native and non-native speakers* (pp. 31-47). Springer.

Wu, T. (2012). *The role of remittances in crisis: An Aceh research study*. Humanitarian Policy Group.

Yuniarni, S. (2016). *Unity in Diversity: Indonesia's Six Largest Ethnic Groups*. Retrieved January 20 from https://jakartaglobe.id/culture/unity-diversity-indonesias-six-largest-ethnic-groups/

Yunus, Y., & Pillai, S. (2020). Pola sebutan monoftong bahasa Jerman oleh penutur bahasa Melayu. *Melayu: Jurnal Antarabangsa Dunia Melayu*, *13*(2), 223-252. https://doi.org/10.37052/jm.13(2)no3

Yusuf, Y. Q. (2013). *A comparative study of vowels in the Acehnese language spoken in Kedah, Malaysia and Aceh, Indonesia* [Doctoral dissertation, Universiti Malaya]. http://studentsrepo.um.edu.my/4459/

Zanten, E. V. (1986). Allophonic variation in the production of Indonesian vowels. *Perception*, *142*, 427-446. https://doi.org/10.1163/22134379-90003348

Zanten, E. V., & Heuven, V. J. v. (1984). The Indonesian vowels as pronounced and perceived by Toba Batak, Sundanese and Javanese speakers. *Bijdragen tot de Taal-, Land- en Volkenkunde*, *140*, 497. https://www.jstor.org/stable/27863619

Zein, M. S. (2017). Elementary English education in Indonesia: Policy developments, current practices, and future prospects: How has Indonesia coped with the demand for teaching English in schools? *English today*, *33*(1), 53-59. https://doi.org/10.1017/S0266078416000407

Zulfadhli. (2015). *A Sociolinguistics Investigation of Acehnese with a focus on West Acehnese: a stigmatised dialect* [Doctoral Degree, University of Adelaide].

Zuraidah, M. D. (1997). Malay+ English= a Malay Variety of English Vowel and Accent. In H. M. Said & K. S. Ng (Eds.), *English is an Asian Language: The Malaysian context* (pp. 35-46).

Zwicker, E., & Terhardt, I. E. (1980). Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *The Journal of the Acoustical Society of America*, *68*(5), 1523-1525. https://doi.org/10.1121/1.385079