AN INSTRUMENTAL STUDY OF FRONT ORAL MONOPHTHONGS IN KENSIU

ADILAH BINTI AZIZAN

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

AN INSTRUMENTAL STUDY OF FRONT ORAL MONOPHTHONGS IN KENSIU

ADILAH BINTI AZIZAN

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS (LINGUISTICS)

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Adilah Binti Azizan

Matric No: TOA170033

Name of Degree: Master of Arts (Linguistics)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

An Instrumental Study of Front Oral Monophthongs In Kensiu

Field of Study: Phonetics/Phonology

I do solemnly and sincerely declare that:

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

Candidate's Signature Date: 15 May 2020

Subscribed and solemnly declared before,

Witness's Signature Date: 15 May 2020

Name:

Designation:

ABSTRACT

This paper acoustically examines the front oral vowels of Kensiu, a highly endangered Austroasiatic language spoken in Peninsular Malaysia. Although there have been initial studies on the phonology of Kensiu, (Bishop, 1996; Burenhult, 2001), these are mainly on the variety spoken in Thailand. This far, there is a lack of studies especially those based on the acoustic analysis. The changes in the sounds of Kensiu are likely to have taken place with increasing contact with the local Malay community in Kedah where the only Kensiu village is located. This paper aims to fill the gap by examining the acoustic properties of front oral monophthongs in Kensiu using a formant frequency model to examine the characteristics of these monophthongs and if there is vowel conflation between these monophthongs. Five female Kensiu native speakers aged between 50 to 62 years old from Kampung Siong in Lubuk Lengong located in Baling, Kedah were recorded. Data were obtained by showing the speakers pictures containing words with the target vowels. Based on measurements of the first formant (F1) and second formant frequencies (F2), the findings suggest that there may be only four front oral vowel phonemes in the Kensiu spoken in Kedah compared to five in the variety spoken in Yala, Thailand. This is due to the conflations of neighbouring vowels as well as possible influence from Malay. This study contributes to our understanding of the Kensiu sound system.

Keywords: Kensiu, indigenous languages, acoustic analysis, monophthongs, formant frequencies

ABSTRAK

Kajian ini secara akustik meneliti vokal oral depan Kensiu, sebuah bahasa Austroasiatik yang sangat terancam yang dituturkan di Semenanjung Malaysia. Walaupun terdapat kajian awal mengenai fonologi Kensiu, (Bishop, 1996; Burenhult, 2001), kajian-kajian ini berfokus pada variasi yang dituturkan di Thailand. Setakat ini, terdapat kelompongan kajian terutamanya yang berdasarkan analisis akustik. Perubahan dalam bunyi Kensiu mungkin berlaku dengan pertembungan dengan masyarakat Melayu di Kedah, di mana satu-satunya penempatan Kensiu terletak. Kajian ini bertujuan untuk mengisi jurang kajian dengan mengkaji ciri-ciri akustik monoftong oral depan dalam Kensiu dengan menggunakan model frekuensi forman untuk mengkaji ciri-ciri monoftong ini dan jika terdapat penggabungan vokal antara monoftong. Lima orang penutur wanita asli Kensiu berumur antara 50 hingga 62 tahun dari Kampung Siong di Lubok Lengong, Baling, Kedah telah direkodkan. Data diperolehi dengan menunjukkan gambar yang mengandungi perkataan dengan vokal sasaran kepada penutur. Berdasarkan ukuran frekuensi forman pertama (F1) dan frekuensi forman forman kedua (F2), dapatan kajian ini mencadangkan bahawa hanya terdapat empat fonem vokal oral depan dalam Kensiu yang dituturkan di Kedah berbanding dengan lima fonem vokal oral depan dalam variasi yang dituturkan di Yala, Thailand. Ini disebabkan oleh penggabungan vokal berdekatan dan kemungkinan pengaruh dari bahasa Melayu. Kajian ini menyumbang kepada pemahaman mengenai sistem bunyi Kensiu.

Kata kunci: Kensiu, bahasa pribumi, analisis akustik, monoftong, frekuensi forman

ACKNOWLEDGEMENTS

First and foremost, I thank Allah for the wisdom He bestowed upon me, the strength and good health He provided me throughout this entire two-year postgraduate journey. It was definitely a challenging journey learning to be comfortable in my own solitude and stepping out from my comfort zone. It was hardly possible for me to complete my master's journey without the precious support from my parents and supervisor.

This dissertation would not have been completed without the encouragement from my supportive supervisor, Prof. Dr. Stefanie Pillai. I would like to offer my gratitude to her for the supervision, thoughtful comments and recommendations on this dissertation. It was from her that I learnt a lot. Thank you for putting up with me in the office for hours to check on my works, sending me simple text messages to ask about my progress and tolerating a lot of things with me. I was secretly grateful for it. I thank Him for leading me to you and I couldn't have asked for better supervisor.

I would like to extend my thanks to my participants for their willingness to provide the data for my study. Despite the challenges and difficulties during almost three months of collecting the data, I was extremely grateful and blessed to be working with a group of supportive participants from the Kensiu community, especially the participants and the Tok Batin, Encik Razali Bin Kulim. It was certainly such an eyeopening experience for me as I had no fieldwork previously.

Last but not the least, the most deserving acknowledgement goes to my mother, Jamilah Binti Senawi, and my father, Azizan Bin Abdul Aziz, for their unending support and prayers. Thank you for being my pillar of strength, for supporting me financially and spiritually, and for the love and care through the phone calls every single day. Thank you for the endless amount of love. For that, I would like to dedicate this dissertation to my parents.

TABLE OF CONTENTS

LIST OF FIGURES

 LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

OA : Orang Asli

JAKOA : Jabatan Kemajuan Orang Asli

UNESCO United Nations Educational, Scientific and Cultural Organization

F1 : Formant 1

F2 : Formant 2

P1 : Participant 1

P2 : Participant 2

P3 : Participant 3

P4 : Participant 4

P5 : Participant 5

ANOVA : Analysis of variation

KM : Kedah Malay

SM : Standard Malay

LIST OF APPENDICES

CHAPTER 1
INTRODUCTION

1.1 Introduction

This chapter begins by providing an overview of the background of the study. It then proceeds to outline the problem statement and research gap. This is followed by the description of the research purposes, research objectives, research questions, significance of the study, and organization of the dissertation.

1.2 Background of the study

The ethnic composition of Malaysia consists of three major ethnic groups; Malay, Chinese and Indian. Malays are the predominant ethnic group which represents 67.4% of the Malaysian population, and the Chinese and Indians comprise 24.6% and 7.3% respectively (Department of Statistics Malaysia, 2015). Malaysia also has an indigenous population. The indigenous people in Peninsular Malaysia are known as *Orang Asli*, while in Sabah they are referred to as *Anak Negeri* and in Sarawak as *Orang Ulu* (Wessendorf, 2008, p. 296). The direct translation of the Malay term Orang Asli (OA) is "original people" and "first people" (Masron, Masami & Norhasimah Ismail, 2013, p. 77). In East Malaysia, the Iban and Bidayuh are the biggest ethnic groups of *Orang Ulu* (also known as *Dayaks*) in Sarawak. In Sabah, the three biggest groups of indigenous people or *Anak Negeri* are *Kadazan-Dusun*, *Bajau* and *Murut*.

Table 1.1 shows the different indigenous groups in Sabah and Sarawak.

Table 1.1. Classification of *Anak Negeri* in Sabah and Orang Ulu in Sarawak

Anak Negeri	Orang Ulu	
Kadazan-Dusun	Iban	Bisayah
Bajau	Bidayuh	Kelabit
Murut	Kenyah	Berawan
Paitan	Kayan	Kejaman
Others	Kedayan	Ukit
	Lunbawang	Sekapan
	Punan	Melana
	Penan	

(Information from the Cultural Survival website https://www.culturalsurvival.org/sites/default/files/Malaysia%202018%20UPR%20Report.pdf)

According to the population statistics of indigenous peoples provided by the Department of Orang Asli Development (2018), there are 178,197 OAs located in Peninsular Malaysia. The department divides the OA population into three subgroups based on cultural and linguistic characteristics: Senoi, Aboriginal or Proto-Malay and Negrito (see Table 1.2 for the categories of OAs in Peninsular Malaysia). There are six tribes in each group. However, smaller groups are sometimes put into the same category as a bigger group for administrative purposes. This is the case with Temoq which comes under Semelai. As can be seen in Table 1.3, the Senoi group has the largest population, followed by Aboriginal Malay. The Negritos comprise the smallest population of OA.

Table 1.2. Categories of Orang Asli in peninsular Malaysia

Senoi	Aboriginal Malay	Negrito
Temiar	Orang Selatar	Kintak
Semai	Jakun	Lanoh
Mah Meri	Kuala	Kensiu
Che Wong	Kanaq	Jahai
Jah Hut	Temuan	Mendriq
Semoq Beri	Semelai	Batek

(Information is taken from the JAKOA website http://www.jakoa.gov.my/)
*The Temoq are put together with a neighbouring group of Semelai for administrative purposed (Laird, 1979)

As shown in Table 1.3, the highest number of OAs live in Pahang followed by Perak, then Selangor, Kelantan, Johor, Negeri Sembilan, Melaka, Terengganu, and the least number of OAs are found in Kedah. The Senois and Negritos are mostly settled in Perak while the majority of Aboriginal Malays live in Pahang. Based on data from data.gov.my, there is no record of OAs residing in Perlis, and Penang. Table 1.3 shows the number of OAs according to states.

Table 1.3. Statistics of OA population according to states

State	Senoi	Aboriginal Malay	Negrito	TOTAL
Johor	55	13,083	1	13,139
Kedah	19		251	270
Kelantan	12,047	29	1381	13,457
Melaka	28	1,486	1	1,515
Negeri Sembilan	96	10,435		10,531
Pahang	29,439	37,142	925	67,506
Perak	50,281	605	2,413	53,299
Selangor	5,073	12,511	3	17,587
Terengganu	818	41	34	893
TOTAL	97,856	75,332	5,009	178,197

(Statistics are taken from the website http://www.data.gov.my/data/ms_MY/dataset/statistik-pendudukmasyarakat-orang-asli-mengikut-sub-etnik)

OA languages in Peninsular Malaysia are made up of two main groups, Austronesian and Austroasiatic (Rohani Mohd Yusof & Nur Hidayah Mohamed Suleiman, 2014). In terms of language, Benjamin (2012a) divides them from a different perspective compared to the classification provided by JAKOA as shown in Figure 1.1, which is based on the genetic relationships of the languages. Benjamin (2012a) considers four Austronesian OA languages to be Malayic dialects: Temuan, Jakun, Kanaq, and Seletar. One other language, Duano, is considered as "an unclassified Austronesian language" (Benjamin, 2012a, p. 141).

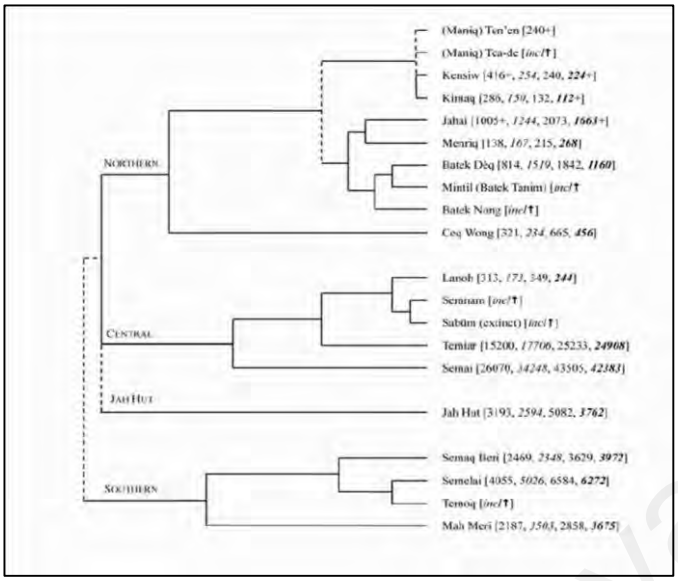


Figure 1.1 Genetic relationships of Aslian languages (Benjamin, 2012a, p. 141)

Based on Benjamin's classification (2012a), the northern OA languages are the one spoken generally by the Negritos, such as Kensiu, Kintaq, Jahai, Menriq and Batek, except for Che Wong. This does not match the JAKOA classification of the OA groups as Che Wong is not under this category. Instead it is classified under the Senoi group by JAKOA. According to Burenhult, Kruspe and Dunn (2011, p. 258) Che Wong is "a geographical outlier of Northern Aslian, spoken by a group of about 300 individuals who are not classified as Semang and whose subsistence is not focused on foraging" which is typically associated with the Northern Aslian groups.

A similar situation applies to the central OA languages where Lanoh is placed together with the languages in the central region along with other Senoic groups, such as Temiar and Semai. Jah Hut is singled out from the central branch and is classified on its own. Dunn, Burenhult, Kruspe, Tufvesson and Becker (2011, p. 295) point out that "Jah Hut is traditionally considered as Central Aslian but difficult to classify with certainty, is spoken by people with mixed societal traditions difficult to assign to the proposed cultural categories". Diffloth and Zide (1992, as cited in Dunn et al., 2011), posit that Jah Hut is the fourth branch of Aslian languages based on the patterns of vowel change and claim

that Jah Hut has a close relationship with Southern Aslian. In relation to this, Benjamin (1976, as discussed in Dunn et al., 2011) raised the question of whether Jah Hut should be classified as a Central Aslian language or if it should be classified separately on its own. As mentioned by Dunn et al. (2011, p. 312), "Jah Hut, Central Aslian and Northern Aslian are distinguished on the same phylogenetic level but remain agnostic about the finer details of subgrouping". The Bayesian phylogenetic inference, which "allow(s) rich inferences from lexical cognate data, by modelling the evolution of a language family as the gain and loss of reflexes of cognate sets" (Burenhult, Kruspe & Dunn, p. 267), is shown in Figure 1.2.

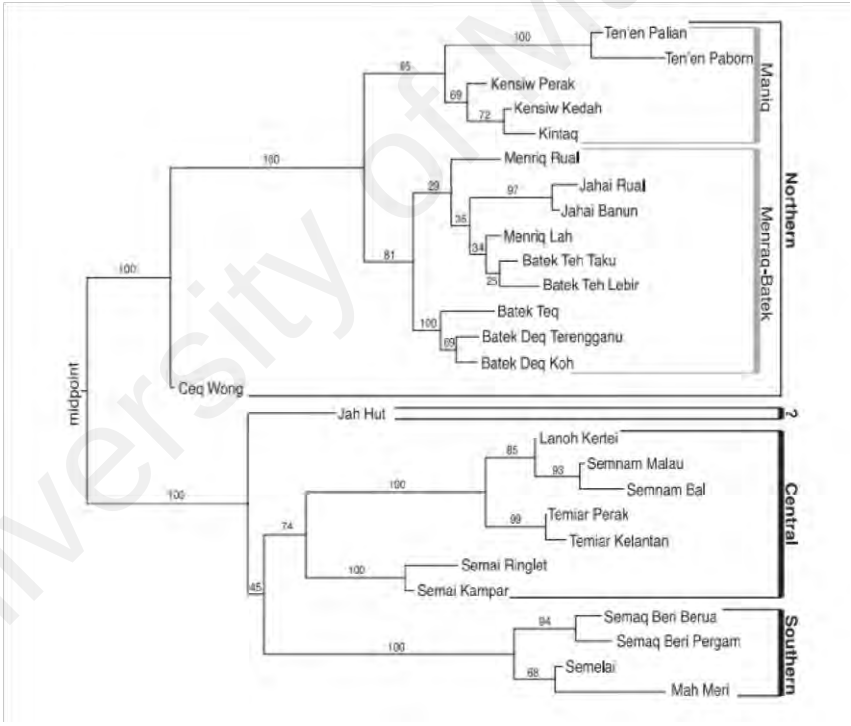


Figure 1.2. Bayesian phylogenetic tree (Dunn et al., 2011, p. 307)

In the southern Aslian languages, it can be seen that there are two Senoic groups which are classified by Benjamin (2012a) under the southern part which is Mah Meri and Semaq Beri. Benjamin (1976) considers Mah Meri to be geographically isolated as they

have higher contact with the Austronesian languages. Thus, that explains why Mah Meri is categorised as a southern OA language rather than a central OA language. In

Benjamin's classification, the Semelais and Temoqs are categorized as different groups. However, as mentioned previously (see Table 1.2), the Temoqs are classified together with their neighbouring group, Semelai. Thus, it can be concluded that the classification of OA groups from JAKOA differs from the classification done by Benjamin (2012a). Dunn et al., (2011) raised an issue that certain Aslian groups are hard to categorise based on their societal-economic segregation due to the reason that some of them may share the same societal-economic features.

Figure 1.3 shows the historical locations of OA languages spoken in Peninsular Malaysia. The map corresponds to the location of each group of OA. Aboriginal Malays comprising the Jakun, Orang Kanaq, Orang Seletar, and Duano who are also known as

"Orang Kuala or Dossin Dolak" (Rohani Mohd Yusof and Nur Hidayah Mohamed Suleiman, 2014, p. 47), are mostly settled in Johor in the southern part of Malaysia. This is in comparison to the Senois who are concentrated in the central part of Malaysia while the northern part of Malaysia is a home for the Negritos including the Kensius.

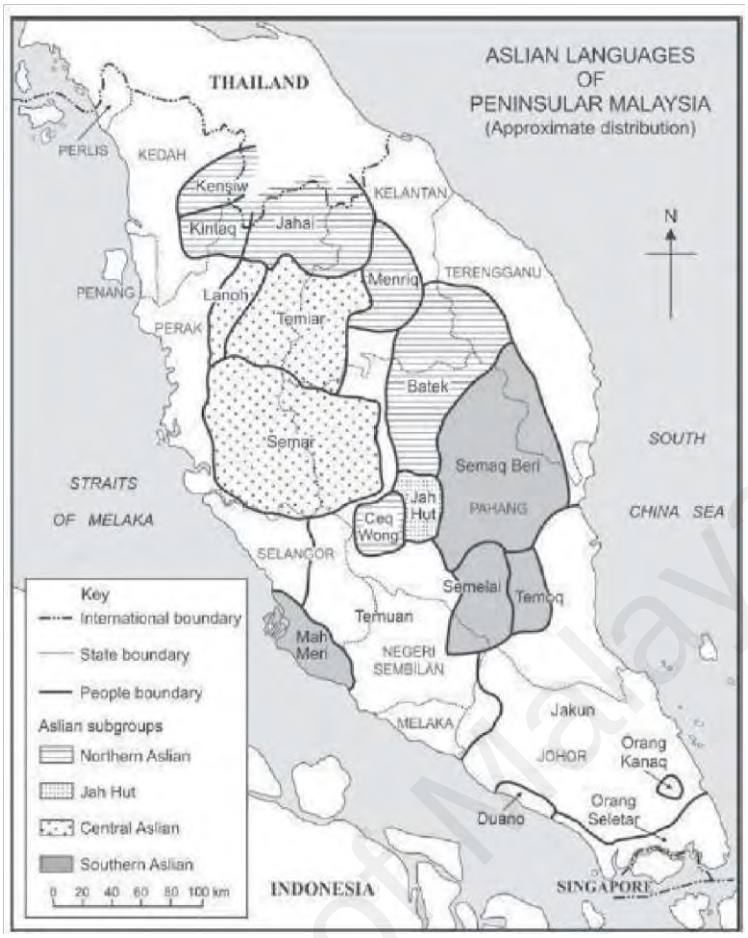


Figure 1.3. Historical distribution of Orang Asli Languages in Peninsular Malaysia (Benjamin, 2012a, p. 144)

1.3 Kensiu people

Kensiu is one of the Austroasiatic groups located in the northern part of Malaysia. According to Simons and Fennig (2017), Kensiu is also known as Kenseu, Kensiw, Mendi, Monik, Moniq, Ngok Pa, Orang Bukit, and Orang Liar. From personal experience with the Kensiu community in Kedah, they sometimes referred themselves to as Maniq. As discussed by Nagata (2005), there are several researchers (e.g. Porath, 2002; Albrecht and Moser, 1998; Bishop, 1996; Hamilton, 2002) who use terms such as Meniq, Mani or Maniq in their studies on the Thai Negritos. However, Nagata (2005) made a remark that these terms have not yet received the status of general acceptance.

The majority of Kensius live in Baling, Kedah and Kampung Siong is their only settlement. However, according to statistics provided by JAKOA (2018), a small

population of Kensiu, 28 people, can be found living in Perak. The Kensius in Kedah are closely related to the Kensius in Yala, Thailand, as geographically the area in which they are located extends into the southern part of the Yala Province (Grimes, 2000).

Kensiu is a northern Aslian language spoken in Kedah and Perak and the language spoken by Kensiu is a Mon-Khmer language which resembles Mon-Khmer languages spoken in Vietnam, Khmer and Munda languages in India (Masron, Masami

& Norhasimah Ismail, 2013). There are said to be several dialects in Kensiu such as Bong, Ijoh (Ijok), Jarum, Jeher (Sakai Tanjong of Temengoh), Kedah (Quedah), Kensiu Batu (Batuq), Kensiu Siong (Siong), Kentaq Nakil (Nakil), Maniq, Plus, and Ulu Selama (Simons & Fennig, 2017). However, there does not appear to be any documented evidence of these dialects. Most of the Kensius can speak at least two languages. Besides their native language, they are also able to speak Malay and a small number of them can understand Thai as well (Carey, 1970). It has been essential for Kensiu to acquire and master another language such as Malay to communicate with people outside their communities. Based on interviews with the Kensius in the present study, they mentioned that they can understand Kintaq as well. In addition, some of them too can understand and speak Temiar (Carey, 1970). They also appear to be able to speak not just the Kedah Malay dialect but also a variety which is similar to the spoken in the central and southern states of Peninsular Malaysia when they speak to people who are not from Kedah. Their linguistic skills are attested by Carey (1970, p.

143): “The Negritos of Baling are, incidentally, excellent linguists; and this despite the fact that practically all of them are illiterate”.

Kensiu is a language in trouble as the number of speakers has been decreasing each year. Fazrul Azmin Zakaria (2010, as cited in Alias Abd Ghani and Salasiah Che

Lah, 2015) stated that Kensiu is at 2.25 on a scale of the United Nations Educational, Scientific and Cultural Organization (UNESCO) level of endangerment, thus making it a severely endangered language under UNESCO's classification. Grimes (2000) stated that there were 3,000 Kensius in Malaysia and Hogan (1984 as cited in Grimes, 2000) claimed there were 3,300 Kensius in Malaysia and Thailand. However, the statistics from JAKOA (2018) indicate that there are 237 Kensius residing in Malaysia. From these statistics, it can be surmised that the population of Kensiu has been decreasing at an alarming rate in 20 years. Apart from the decreasing numbers, the younger generations of Kensiu tend to favor Kedah Malay compared to their native language, meaning that the number of Kensiu speakers has also been declining. Hence, as time passes by, Kensiu is rarely spoken among the younger generations. It is possible that in the future, Kensiu will be on the list of extinct languages.

1.4 Problem statement

Previously published studies Malaysian Kensiu have mainly been sociolinguistic in nature (e.g., Alias Abd Ghani and Salasiah Che Lah, 2015; Roshidah Hassan, Kamila Ghazali & Asmah Haji Omar, 2015). The study on the phonological aspect of Kensiu was solely based on the data collected in Yala in southern Thailand, for example, the study by Bishop (1996). Burenhults' (2001) study was on the linguistic aspects of the Negrito groups in Thailand and Malaysia including the Kensiu spoken in Malaysia and Thailand. The sound system of the variety spoken in Malaysia remains understudied.

1.5 Research gap

As mentioned in the previous section, whilst some studies have been carried out on Kensiu, there is a lack of published studies on the phonemes in Kensiu in Malaysia. Further, these studies were based upon data from over 20 years ago, and there have been no published studies on Kensiu sounds since then. Therefore, there is a gap in our knowledge of Kensiu sounds, and the present study aims to begin to fill this gap by acoustically examining vowels in Kensiu, specifically its front vowels.

1.6 Research purpose, research objectives, and research questions

In order to address part of the research gap, the overall aim of this study is to determine the characteristics of the front oral monophthongs in Kensiu by examining their acoustic properties. In relation to the research purpose, the main objectives of this study are:

1) to determine characteristics of the front oral monophthongs in Kensiu based on first (F1) and second formant frequencies (F2) measurements of the vowels.

2) to determine if there is vowel conflation between these front oral monophthongs based on their first (F1) and second formant frequency (F2) measurements. In relation to these objectives, this study seeks to address the following research questions:-

1) What are the characteristics of the front oral monophthongs in Kensiu based on the first (F1) and second formant frequency (F2) measurements of the vowels?

2) To what extent is there vowel conflation between these front oral monophthongs?

1.7 Significance of the study

There are numerous languages that are rarely given any recognition because people who speak those languages are a minority community, and considered to be socially and economically disadvantaged with no political 'value'. Therefore, the languages and cultures which are part of their inheritance and identity are often doomed to extinction. Further, globalization is one of the threats to languages and cultures which are already vulnerable. This study offers insights into the Kensiu sound system through an acoustic study and has the potential to add to previous studies on the sound system of Kensiu. Prior to this study, it was difficult to make predictions about the production of

vowels in Kensiu as the latest study of Kensiu phonology was about 20 years ago with the detailed research being on the Thai variety of Kensiu (Bishop, 1996). The findings reported in this study can provide insights into the current sound system of Kensiu, specifically its vowels. It is hoped that this study can also contribute to the Kensiu community itself as one of the ways to document an aspect of their language.

1.8 Organization of the dissertation

The dissertation has been organized into five themed chapters. This study begins by explaining the background of research in Chapter 1. It will then go on to explain to the problem statement, research gap, research purpose, research objectives, research questions, and significance of the study. Chapter 2 begins by discussing previous studies on Kensiu and other OA languages. It also examines how vowels are acoustically analysed. The third chapter is concerned with the methodology used in this study. Chapter 4 presents and discusses the findings of the study. Lastly, the findings based on the research questions will be summarized in Chapter 5, which also presents the limitations of the study and the recommendations for future studies.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

This chapter begins with an explanation of Austroasiatic and Austronesian languages. The chapter then moves on to the language endangerment and OA languages, and Malay borrowings in OA languages. Then, it proceeds to discuss previous studies on OA sounds and followed by Kensiu sounds. This chapter also addresses Kedah Malay and Standard Malay in comparison to Kensiu. The final section in this chapter explains the formants measurements and vowels and the studies on the acoustic analysis of vowels are then discussed.

2.2 Austroasiatic and Austronesian languages

Indigenous people in Peninsular Malaysia are categorized under two major language families, the Austroasiatic and Austronesian language families. The Senois and Negritos belong to the Austroasiatic language family while the southern part of OA, Aboriginal Malays comes under the Austronesian language family.

Austroasiatic languages are the language family that is primarily divided into two big branches, Munda languages in India and Mon-Khmer in Southeast Asia (Reid, 1994). There are many proposals on the Austroasiatic dispersal. However, there is a lack of evidence to support those claims (Sidwell, 2009). Austroasiatic families are said to have dispersed from South China in Yunnan Province, Kampuchea, Malaysia, and

Vietnam in the eastern part, then from Central and Eastern India to Nicobar Islands, Burma, Thailand, and Laos (Reid, 1994). Table 2.1 presents a classification of

Austroasiatic languages from Sidwell (2009).

Table 2.1. Branches of Austroasiatic language (Sidwell, 2009, p. 3)

Branch	Regions spoken
Aslian	Malay Peninsular
Bahnaric	Central Indochina
Katuic	Central Indochina
Khasic	Maghalaya State of India
Khmer	Cambodia and neighbouring areas
Khmunic	Northern Laos
Monic	Southern Myanmar and central Thailand
Munda	Eastern and Central India
Nicobaric	Nicobar Islands of India
Palaungic	Shan State of Myanmar
Pearic	Cambodia and Thailand
Vietic	Vietnam and Central Laos

Austronesian speaking communities can be found in four main parts of Southeast Asia (Blust, 2013):

a) The Malay Peninsular

b) Coastal islands of peninsular Burma and Thailand

c) Interior regions of Vietnam, Laos, and Cambodia

d) Hainan Island in southern China

As mentioned in Chapter 1, there are several OA Austronesian languages. Blust (2013) illustrates ten of the largest Austronesian languages in Southeast Asia as shown in Figure 2.1.

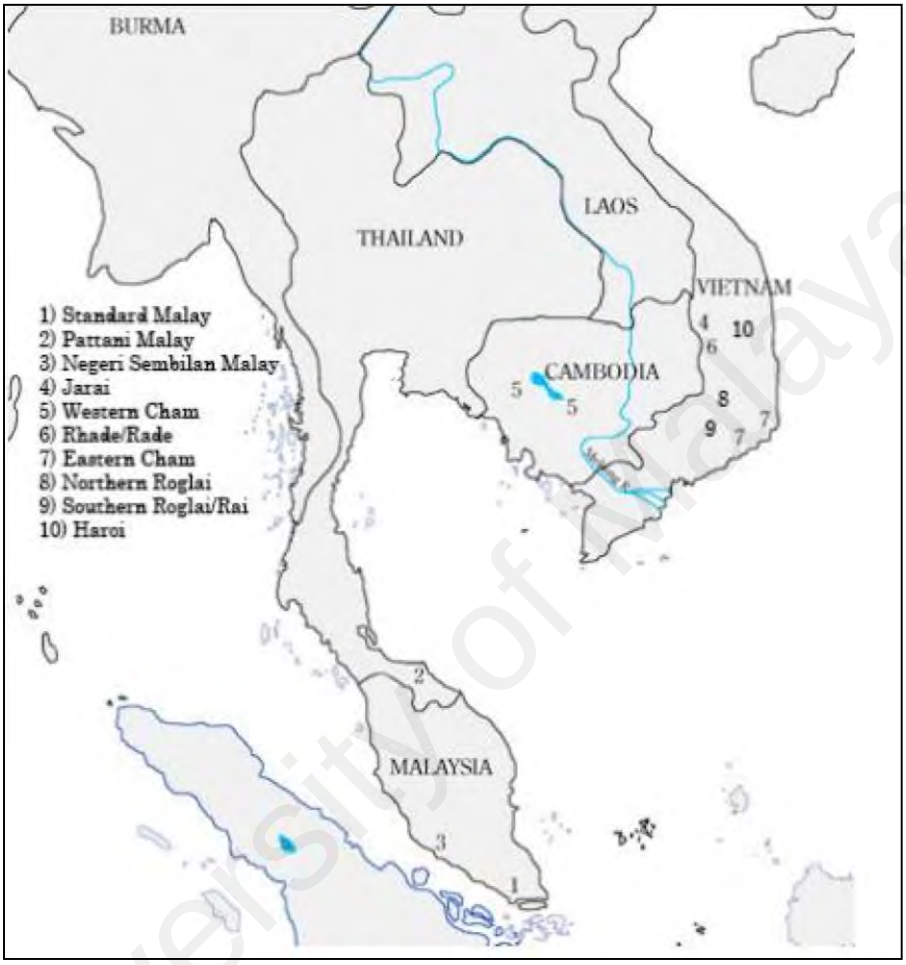


Figure 2.1. The largest Austronesian languages in Southeast Asia (Blust, 2013, p. 72)

2.3 Language endangerment and OA languages

Crawford (1995, as cited in Benjamin, 2012b) explains that language loss theories can be seen in two ways, "murder" and "suicide". Language murder and language suicide occur in different circumstances. The former is said to occur "when political pressures against the language become too great to withstand". An example of this is the institutionalised attempt at wiping out Native American languages in the 1860s

(Crawford, 1995). Such is the case with Upper Necaxa Totonac, in East-Central Mexico which has been taken over by Spanish (Beck & Lam, 2008).

Lewis and Simons (2010, as cited in Coluzzi, 2017) introduce EGIDS which has ten levels of language vitality. EGIDS is the expansion of GIDS (Graded Intergenerational Disruption Scale) created by Fishman (1991, as cited in Coluzzi, 2017). The descriptions of EGIDS are more general as it excludes the number the speakers, language attitudes, language policies and language documentation. Table 2.2 shows ten levels of EGIDS.

Table 2.2. EGIDS (Lewis & Simons, 2010, as cited in Coluzzi, 2017, p. 213)

Level	Label	Description
0	International	The language is used internationally for various purposes.
1	National	The language is used at nationwide level such as in education, work, mass media, and government.
2	Regional	The language is used for local and regional mass media and government services.
3	Trade	The language is used for local and regional work by both insiders and outsiders.
4	Educational	The language is used in a public education.
5	Written	The language is used orally as well in a written form
6a	Vigorous	The language is used orally by all generations and is being learned by children as their first language.
6b	Threatened	The language is orally by all the generations but only some parents' generations transmitting the language to their children.
7	Shifting	The parent generation knows the languages but none are transferring it to their children.
8a	Moribund	The language is used only by the grandparent generation as the active speakers.
8b	Nearly extinct	The language is used by the grandparent generation or older who have little opportunity to speak the language.

9	Dormant	The language functions as an identity for an ethnic community. No one has more than symbolic proficiency.
10	Extinct	No one maintains the identity of the language.

Duangchan (2006) put forward four critical factors that lead to language endangerment especially in northern Aslian languages in Thailand such as Kansiw, Teade, Yahay, and Tean-ean. These languages are spoken in five provinces which are Yala Province, Narathiwat Province, Satun Province, Trang Province, and Phattahalung Province. The first factor that contributes to language endangerment is frequent contact with the outsiders requiring them to learn and speak dominant languages such as Thai and Malay. The speakers may also feel ashamed to speak their own native language as it is only spoken by the minorities. Second, the languages may not have an orthographic system. Third, the government restricts their settlement areas, and this leads to them practising the languages and cultures of the dominant groups. Lastly, modernization has swayed their traditional way of life.

In fact, language shift among OAs is also one of the factors contributing to language endangerment in these minority languages. In Malaysia, *Bahasa Malaysia* is a national language that is widely used as the medium of instruction in government schools and the local geographical Malay dialect is generally also the lingua franca in the states that the OAs live in. In fact, at present, only three indigenous languages are taught in schools such as Iban, Kadazan and Semai. Thus, Renganathan and Kral (2018) claimed that indigenous languages such as OA languages are perceived to be less prestigious as it is only spoken at home and among their community compared to the dominant languages such as Malay and English which are seen as economically

advantages as it allows the speakers to have better future.

As an example, the use of Mah Meri, an OA language, is slowly shifting towards Malay especially among the younger generation (Coluzzi, Riget & Xiaomei, 2017). The Mah Meri youth regularly adapt words and sentence structures from Malay. However, although Malay is slowly becoming emerging into Mah Meri, the findings of Coluzzi,

Riget, and Xiaomei (2017) revealed that Mah Meri still appears to be a dominant language among them. It was reported that 93% of the respondents mentioned Mah Meri as the most fluent language they speak and only 4.7% claimed Malay as their most fluent language (Coluzzi, Riget, Xiaomei, 2017, p. 142). However, there was a clear pattern of language shift from Mah Meri in the family domain. It was reported that the percentage of the use of Mah Meri decreased by almost half from 87.2% with grandparents to 81.4% with parents, and 47.7% with the children (Coluzzi, Riget, Xiaomei, 2017, p. 142). Language shift has also permeated other domains, and with the domain of family being affected, the language becomes endangered (Heinrich, Miyara & Shimoji, 2015).

A similar thing is happening with Kensiu. As discussed in Chapter 1 (see 1.3), the younger generation of Kensius tend to favour Kedah Malay over their native language. Thus, the language is rarely spoken among the youth. However, as time passes by, the older generations also seem to have shifted their language to Kedah Malay. This can be seen in the study by Alias Abd Ghani and Salasiah Che Lah (2015), where their findings revealed that the majority of the elderly Kensiu speakers appear to have shifted their language to Malay.

2.4 Malay borrowings in OA languages

Hockett (1985, as cited in Hoffer, 2005) introduced four processes of language borrowings: loanwords, loan shifts, loan-translations and loan-blends. Loanwords usually occur in OA languages where the speakers adopt the word from the source language such as Malay. The frequency of Malay loanwords in OA is high and common due to the contact during Malay migration (Tengku Intan Suzila & Teo, 2015).

In the case of Kensiu, the Kensius consider their native language as the language of intimacy and kinship, and thus, they generally speak their language among themselves

(Alias Abd Ghani & Salasiah Che Lah, 2015). However, Alias Abd Ghani and Salasiah Che Lah (2015, p. 25) found that, 67% of the lexical items in Kensiu were Malay loanwords (see Table 2.3), and there were a considerable number of Kensiu words which were code-mixed with Malay (see Table 2.4).

Table 2.3. The Malay loan words (Alias Abd Ghani & Salasiah Che Lah, 2015, p. 26)

Malay	Semang Kensiu
nasi (rice)	nasi
dapur (kitchen)	dapur
pintu (door)	pintu
tingkap (window)	tingkat
kucing (cat)	kucing
lembu (cow)	lembu
bas (bus)	bas
lori (lorry)	lori
kasut (shoes)	kasut
kampung (village)	kampong
bandar (town)	bandar

Table 2.4. The Semang Kensiu-Malay code-mixed words (Alias Abd Ghani & Salasiah Che Lah, 2015, p. 26)

Malay	Semang Kensiu
kutip buah (collecting fruits)	kutip kebek
kumpul rotan (collecting rattan)	kumpul awei
pokok bunga (flowering plants)	tom bungak
makan pagi (breakfast)	cik pagik
makan malam (dinner)	cik malam
daun kesum (kesum leaves)	helik kesum
daun selasih (selasih leaves)	helik selasih
batang pokok (tree trunk)	batang ihuk
bilik air (toilet)	bilik betew

In addition, Alias Abd Ghani and Salasiah Che Lah (2015) also provide the borrowings of Malay lexical items to Kensiu that are mostly related to modernization as shown in Table 2.5.

Table 2.5. Semang Kensiu words associated with modern living and technologies (Alias Abd Ghani & Salasiah Che Lah, 2015, p. 26-27)

Malay	Semang Kensiu
televisyen (television)	tv
hospital (hospital)	hospital
kelinik (clinic)	kelinik
doktor (doctor)	doktor
pensil (pencil)	pensil
radio (radio)	radio
bas (bus)	bas
komputer (computer)	komputer
kapal terbang (aeroplane)	kapel terbang
telefon bimbit (mobile phone)	telefon bimbit

The influence of Malay words can also be seen in Duangchan's (2006) study of four northern Aslian languages. A previous study by Bishop (1996) also claimed that the analysis in her study was uncertain due to a large number of Malay borrowings and the small number of people in Kensiu. Duangchan (2006) presents 472 vocabulary items of Kansiw (Duangchan's spelling of Kensiu), Tae-de, Yahay and Tean-ean and Bishop (1996) presents 2,170 items in the Kensiw (Bishop's spelling of Kensiu) glossary. Through this list, it can be seen that there are a substantial number of loan words from Malay that have been used in. As Duangchan (2006, p. 208) points out,"(t)hese four northern Aslian languages use loan words from other languages, especially Malay". This is supported by Yager and Burenhult (2017) that Aslian speakers usually speak three or more languages excellently, such as Malay and Thai. Table 2.6 shows some of the Malay loan words in Kensiu in Duangchan's (2006) study.

Table 2.6. Malay loan words in Kansiw (Duangchan, 2006, p. 210-216)

Kansiw (Duangchan, 2006)	Malay	English (Duangchan, 2006)
santan	santan	coconut
blacɛn	belacan	shrimp paste
cukko	cukur	to shave off
sampah	sampah	garbage
laŋkah	langkah	to cross
malɛs	malas	lazy
dapo	dapur	kitchen
sampit	sempit	narrow
tabaŋ	tebang	to cut down
buyaʔ	buaya	crocodile
pucat	pucat	pale
kacaŋ	kacang	beans

*Note: There is no difference between Kensiw, and Kansiw. Kensiw (as used by Bishop, 1996) or Kansiw (as used by Duangchan, 2006) generally efers to a group that lives in the Yala Province, in Thailand.

2.5 Previous studies on OA sounds

In general, Aslian languages present a vowel system that consists of three to five levels of vowel heights (Burenhult, 2001). This is similar to Kruspe (2004) who says that there are three levels of vowel height and three levels of vowel frontness in all Aslian languages. Kruspe (2004) also suggests that there are also phonemic contrasts of nasalisation in these languages. Table 2.7 and Table 2.8 provide the standard Aslian vowels inventory for both oral and nasal vowels.

Table 2.7. Standard Aslian oral vowels (Kruspe, 2004, p. 59)

Oral

i	ʉ	u
e	ə	o
ɛ	a	ɔ

Table 2.8. Standard Aslian nasal vowels (Kruspe, 2004, p. 59)

Nasal		
ĩ	ʉ̃	ũ
ẽ	ə̃	õ
ɛ̃	ã	ɔ̃

In the consonant inventory of Aslian languages, there are five places of articulation and six manners of articulation as mentioned by Benjamin (1985, as cited in Kruspe, 2004). Table 2.9 shows the consonant inventory of Aslian languages.

Table 2.9. Consonants inventory (Benjamin, 1985, p. 58) (as cited in Kruspe, 2004)

	Bilabial		**alveolar**		**palatal**		**velar**		**glottal**
Stop (aspirated)	p p^h	B	t t^h	d	c c^h	j	k k^h	g	ʔ
Nasals (+ glottal)	(m)	m ʔm	(n)	n ʔn		ɲ ʔɲ	(ŋ)	ŋ	
fricatives	(ɸ)				s	(z)			h
laterals (+glottal)				l (ʔl)					
rhotics (+glottal)				r (ʔr)					
Glides (+glottal)						w		y (ʔy)	

From Table 2.9, it can be seen that generally in Aslian languages, there are four aspirated stops in four places of articulation except for glottal. In comparison to Bishop

(1996), there are three aspirated stops such /kʰ/, /pʰ/ and /tʰ/ in the Kensiu spoken in Yala, Thailand. These three aspirations are the result of borrowed words from Malay and Thai, and they do not frequently occur (Bishop, 1996). The /kʰ/ and /tʰ/ are the result of borrowings from Thai, while /pʰ/ occurs in Malay borrowings. These three aspirated stops appear in syllable-initial positions (Bishop, 1996).

Since Austroasiatic languages have complicated vowel systems, Benjamin (2012b), says that it is insufficient to transcribing or hearing these vowels in *a e i o u*. Benjamin (2012b) presented the vowel inventories in Temiar, Jah Hut, Jahai, Kampar Semai, Semelai, Ceq Wong, and Semaq Beri in Table 2.10 and Table 2.11. From these tables, it can be concluded that Temiar, Jah Hut, and Jahai share the same vowel inventory while the vowel distributions in Kampar Semai, Semelai, Ceq Wong and Semaq Beri are the same. There are ten vowels in Temiar, Jah Hut and Jahai and twelve vowels in Kampar Semai, Semelai, Ceq Wong and Semaq Beri.

Table 2.10. Temiar, Jah Hut, Jahai vowels (Benjamin, 2012b, p. 13)

	Front	**Central**	**Back**
High	i	ʉ/ɨ	u
Mid	e	ə	o
Low	ɛ	a	ɔ

Table 2.11. Kampar Semai, Semelai, Ceq Wong and Semaq Beri vowels (Benjamin, 2012b, p. 13)

	Front	**Central**	**Back**
High	i	ɯ/ʉ	u
Mid-high	e	ə	o
Mid-low	ɛ a/æ		ɔ
Low	a/æ		ɒ

More specific studies that focused on the phonological aspect of a particular OA language is Stevens, Kruspe and Hajek (2006) who looked at Mah Meri. This study focused on the phonetic analysis of register in Mah Meri. Most of the Aslian languages

especially Mon Khmer languages do not have prosodic features like voice registers (Stevens, Kruspe & Hajek, 2006; Kruspe, 2004). Stevens, Kruspe, and Hajek (2006) defined voice register as the complex of various laryngeal and supralaryngeal events that related to voice and vowel quality, length and pitch. The difference between register 1 and register 2 lies in with voice quality, duration and pitch. Table 2.12 shows the vowels registers in Stevens, Kruspe, and Hajek (2006, p. 2).

Table 2.12. Vowels register (Stevens, Kruspe & Hajek, 2006, p. 2)

Register 1 vowels	**Front**		**Central**		**Back (-round)**		**Back (+ round)**	
High	i	i̤			ɯ	ɯ̤	u	ṳ
Mid High	e	e̤					o	o̤
Mid Low	ɛ	ɛ̤	ə	ə̤			ɔ	ɔ̤
Low			a	a̤				

The analysis in this study was predominantly based on auditory evaluation from 32 tokens produced by a male speaker. These words were elicited in isolation from a paired token for each vowel. Table 2.13 shows the lexical items for each vowel and register.

Table 2.13. Lexical items for each vowel and register (Stevens, Kruspe & Hajek, 2006, p. 3)

Vowel quality	**Register 1**	**Gloss**	**Register 2**	**Gloss**
a	luwat	‘mangrove worm’	luwa̤t	‘front’
e	ket	‘little’	ʔiʔe̤t	‘no, not’
ɛ	jɛc	‘be bored’	sɛ̤c	‘endpoint’
u	bəkut	‘be blunt’	dṳk	‘house’
o	jok	‘to uproot’	co̤k	‘rattan’

ə	ɯ̟ə̟c	‘Munia sp. bird’	ɯ̟ə̟c̣	‘to throb’
ɔ	sɔp	‘to dress’	k^{h}ɔ̟p	‘to get’
ɯ	bɯt	‘to stop running’	təkɯ̟t	‘to press’

Based on an auditory evaluation, vowels in register 1 were found to have a clear, tense voice quality, shorter duration and lower pitch as opposed to register 2 which had a longer duration and high pitch (Stevens, Kruspe & Hajek, 2006). Register 1 in the spectrogram in Figure 2.2 is clearly more defined compared to register 2. The findings also indicated that there were no significant differences in terms of vowel duration although register 2 displayed a high duration based on the previous impressionistic analysis. Figure 2.2 shows the spectrogram of register 1 and register 2 vowels for the words /luwat/ and / luwạt/.

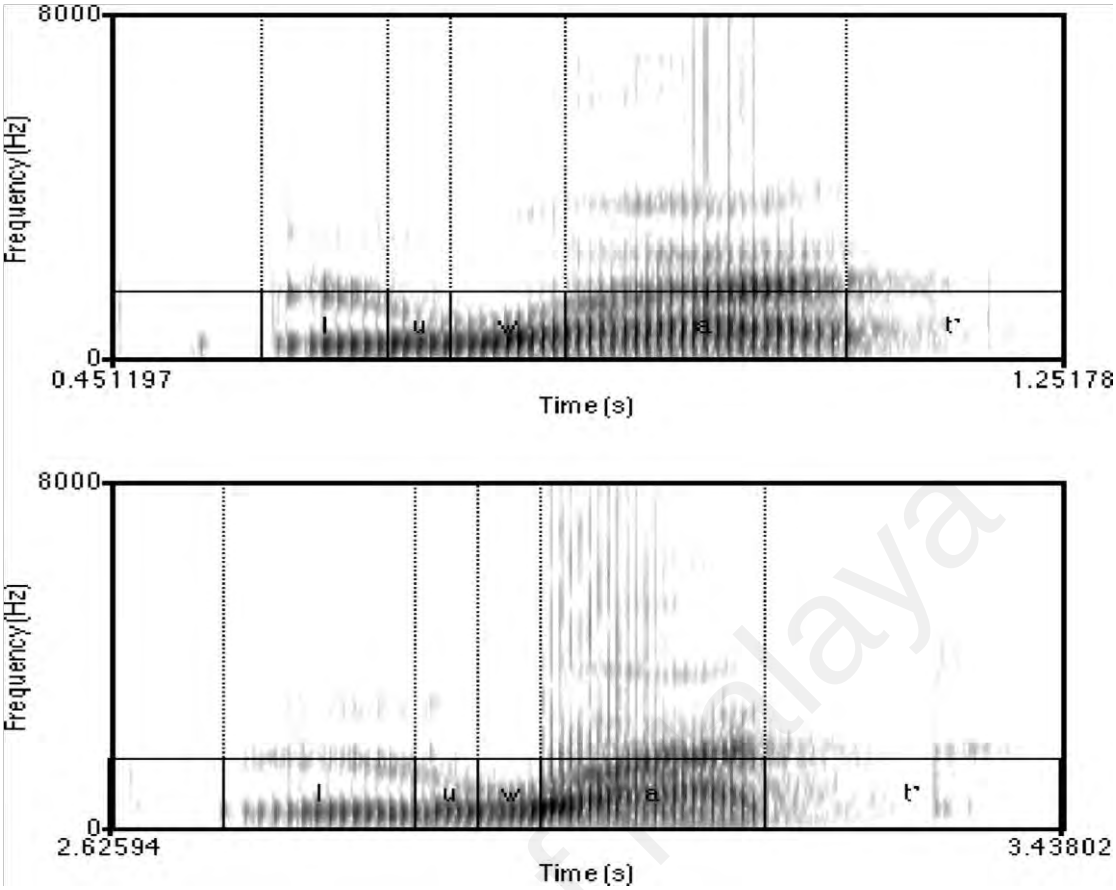


Figure 2.2. Spectrogram of /luwat/ and / luwa̤t/ (Stevens, Kruspe & Hajek, 2006, p. 3)

Phillips's (2007) study mentioned the long and short vowels in Semai. This paper described the phonology of the Betau dialect, one of the dialects of Semai. Phillips (2007) suggested that there are 45 phonemes in the Betau dialect, including 19 consonants, 14 oral vowels and 12 nasal vowels as shown in Table 2.14. The vowel inventory has both oral long and short vowels as well as nasal long and short vowels.

Table 2.14. Vowel inventories of Semai Betau (Phillips, 2007, p. 5)

Oral Vowels, long	**Front (unrounded)**	**Central (unrounded)**	**Back (rounded)**
Close	ii	~~ɨɨ~~	uu
Close-mid	ee		oo
Open-mid	ɛɛ		ɔɔ
Open		aa	
Oral Vowels, short			
Close	i		u
Mid	ɛ	ə	ɔ
Open		a	

Nasal Vowels, long			
Close	ĩĩ̃	ɨɨ̃	ũũ
Mid	ɛɛ̃		ɔɔ̃
Open		aã̃	
Nasal Vowels, short			
Close	ĩ̃		ũ̃
Mid	ɛ̃	ə̃	ɔ̃
Open		ã̃	

Temiar has a complex vowel inventory as well, with short and long oral vowels and nasal vowels as well (Benjamin, 2012b). Table 2.15 shows the vowel inventory for Temiar.

Table 2.15. Vowels inventory in Temiar (Benjamin, 2012b, p. 13)

	Front	**Central**	**Back**
Short oral	i	ɯ	u
	e	ə	o
	ɛ	a	ɔ
Long oral	ii	ɯɯ	uu
	ee	əə	oo
	ɛɛ	aa	ɔɔ
Short nasal	ĩ̃	ɯ̃	ũ̃
	-	-	-
	ɛ̃	ã̃	ɔ̃
Long nasal	ĩĩ̃	ɯ̃ɯ̃	ũũ
	-	-	-
	ɛɛ̃	aã̃	ɔɔ̃

2.6 Kensiu sounds

In the Kensiu variety spoken in Yala Thailand, there are said to be 28 vowels comprising 14 oral monophthongs, 12 nasal monophthongs and two diphthongs (one oral diphthong and one nasal diphthong) (Bishop, 1996). There are five vowel heights in Kensiu which are close, close-mid, mid, open-mid and open and a presence of three vowel fronting, front, central and back. Thus, five vowels are located at the front of the vowel space and four vowels are located at the back vowels (Bishop, 1996). Table 2.16 shows the Kensiu vowels as reported in Bishop (1996).

Table 2.16. Kensiu vowels in Bishop (1996, p. 228)

	Front		**Central**		**Back**	
Tongue Height	oral	nasal	oral	nasal	oral	nasal
Close	i	ĩ	ɯ	ɯ̃	u	ũ
Near-Close	ɪ	ɪ̃				
Close-Mid	ɐ	ɐ̃	ɤ		o˜	o˜
Mid	e	ẽ	ə		o	õ
Open-Mid	ɛ	ɛ˜	ʌ	ʌ̃	ɔ	ɔ̃
Open			a	ã		
Diphthongs	ie	i ẽ˜				

Note: Kensiu is spelt as Kensiw in Bishop (1996)

Kruspe (2004) stated that most of the Aslian languages have no contrastive register and contour tone but in Bishop (1996, p. 238) stated that "Kensiw is not a tonal language, but there is a very small number of pairs of lexeme that contrasts only on the basis of a pitch difference". Bishop (1996) describes the pitch difference as one having a normal pitch and the other as having high pitch. Table 2.17 provides the lexical items in Kensiu that have pitch differences.

Table 2.17. Pitch differences in Kensiu (Bishop, 1996, p. 238)

Normal pitch	High pitch
gūj	gúj
kɛ̄	kɛ́
kā	ká

Note: Kensiu is spelt as Kensiw in Bishop (1996)

Bishop (1996) states that there are two aspirated stops in /pʰ / and /tʰ / in Kensiu which occurs in initial syllable in Malay borrowings such as *paɻo̤* t [pʰaɻo̤t] 'scars' and in Thai borrowings *kata* [katʰaʔ] 'frying pan, pot'. However, as mentioned by Bishop (1996), /p/ and /t/ have rare aspirated allophones, and thus, it could be due to the rare occurrences of /pʰ / and /tʰ/ that these allophones are not listed in Bishop's (1996) consonants inventory. Table 2.18 shows the consonant phonemes in Kensiu from Bishop (1996).

Table 2.18. Consonant phonemes in Kensiu (Bishop, 1996, p. 232)

	bilabial	alveolar	retro	palatal	velar	glottal
stop, vl.	p	t		c	k	ʔ
stop, vl. asp.					kʰ	
stop, vd	b	d		ɟ	g	
nasal	m	n		ɲ	ŋ	
fricative, vl.	ɸ	s				h
fricative, vd.					ɣ	
lateral approximant		l				
central approximant	w		ɻ	j		

Note: Kensiu is spelt as Kensiw in Bishop (1996)
vl.: voiceless, vd.: voiced, asp.:aspirated

2.7 Kedah Malay and Standard Malay

As discussed earlier, it is likely that indigenous languages are very much influenced by Malay. Malaysians from various ethnic groups, whether they are from the

Austronesian family or non-Austronesian family, for several centuries have been using the Malay language as the intermediate language between groups (Asmah Haji Omar, 1991). According to Teoh (1994), Standard Malay has six vowel systems that are divided into high, low and back. The representation of the Standard Malay vowel inventory is illustrated in Table 2.19.

Table 2.19. Standard Malay vowels inventory (Teoh, 1994, p. 12)

	Front	**Central**	**Back**
High	i		u
Mid	e	ə	o
Low		a	

Teoh's (1994) Standard Malay vowel inventory is consistent with Indirawati Zahid and Abdul Hamid Mahmood (2016) except for the vowel /a/. The vowel /a/ in Indirawati Zahid and Abdul Hamid Mahmood (2016) is located at the open and front positions as opposed to Teoh (1988), where it is located at the low and central. Figure 2.3 shows the quadrilateral vowel chart for Standard Malay.

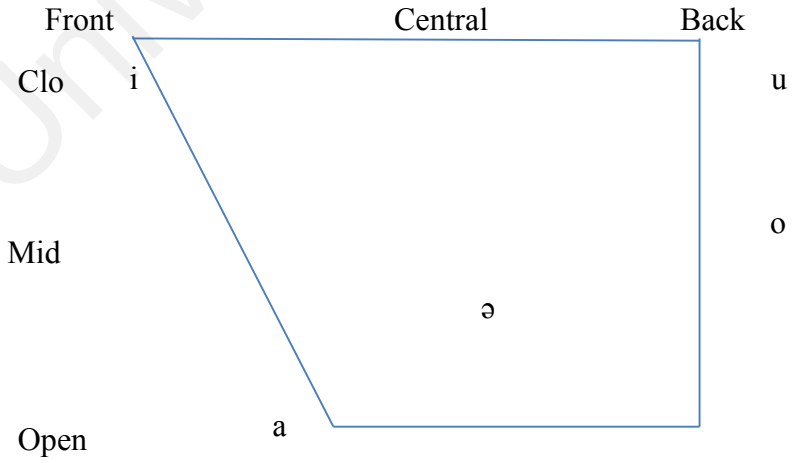


Figure 2.3. Vowel chart of Standard Malay (Indirawati Zahid & Abdul Hamid Mahmood, 2016, p.

The term Standard Malay is derived from one of the dialects spoken commonly in the southern part of peninsular Malaysia, which is the Johor-Riau Malay dialect (Teoh, 1988). The similarity of the vowel systems between standard Malay and the Johor dialect is due to the historical events that led to the Johor dialect becoming the base for the standard language in Malaysia (Asmah Haji Omar, 1991). Apart from Standard Malay and the Johor dialect which have six vowels, other Malay dialects, including Perak, Kedah, Pulau Pinang, Negeri Sembilan, Sarawak, and Kelantan have eight vowels as shown in Table 2.20 (Asmah Haji Omar, 1991).

Table 2.20. Malay dialects vowels (Asmah Haji Omar, 1991, p. 22)

Malay dialects	Vowels
Standard Malay and Johor	/i, e, ə, a, u, o/
Perak, Kedah, Pulau Pinang, Negeri Sembilan, and Sarawak	/i, e, ɛ, ə, a, u, o, ɔ/
Kelantan	/ɨ, ɛ̈, ë, ə, a, ʉ, ö , ɔ̈/

Yusuf and Pillai (2016) discussed some of the distinctive features between standard Malay and Kedah Malay. Table 2.21 shows the distinctive features between standard Malay and Kedah Malay.

Table 2.21. Standard Malay and Kedah Malay words (Yusuf & Pillai, 2016, p. 16)

	Standard Malay	Kedah Malay
/ə/ - /a/	[apə] ‘what’	[apa] ‘what’
/il/ -/e/	[katil] ‘bed’	[kate] ‘bed’

/el/- /ɛ/	[tʃomel] ‘cute’	[tʃomɛ] ‘cute’
/oh/ -/ɔ/	[dʒodoh] ‘fate’	[dʒɔdɔ] ‘fate’

2.8 Formants measurements and vowels

In acoustic studies, formant frequencies are used to identify the formant pattern of vowels (Kent & Read, 2002). Dowd, Smith and Wolfe (1998) described formants as the resonances of the vocal tract in the acoustical phonetics. Kent and Read (2002) introduced the rule of thumb with regard to vowel formant frequencies to vowel articulation where the first format (F1) is related to tongue height and the second format (F2) to tongue advancement. Generally, high vowels have low F1 frequencies and low vowels have high F1 frequencies (Kent & Read, 2002). For F2, front vowels have high F2 frequencies and back vowels have low F2 frequencies (Kent & Read, 2002). In other words, F1 represents the open and close positions and F2 represents the front and back positions to show the difference in the vowels plotting (Ladefoged, 1993 as cited in Deterding, 2003).

The characteristics of vowels can be described using the first formant and second formant only, but in the case of high vowels and r-coloured vowels, the third formant can also be measured (Ladefoged, 2003).

Table 2.22 shows the example of the first two formants F1 and F2 for the five vowels of American English by Hillenbrand et al. (1995) and Assman and Katz (2000) (as cited in Kent & Read, 2002). From this table, it can be seen that the highest F1 is in the vowel /a/, the low vowel and the highest F2 frequency is in the vowel /i/ which is more fronted than the other vowels.

Table 2.22. F1 and F2 measurements of five vowels in American English (Kent & Read, 2002, p. 112)

	I	**e**	**a**	**o**	**u**

F1	437 (1)	429 (2)	536 (1)	572 (2)	936 (1)	688 (2)	555 (1)	636 (2)	459 (1)	430 (1)
F2	2761 (1)	2588 (2)	2530 (1)	2309 (2)	1551 (1)	1273 (2)	1035 (1)	1470 (2)	1105 (1)	1755 (2)

(1) Hillenbrand et al. (1995) (2) Assman and Katz (2000)

This formant frequency model is used in studies to identify each vowel by listening to the recordings and the linear-prediction-based formant tracks will be derived from spectrograms (Deterding, 1997). Acoustic signals are visually shown in the spectrograms. A dark band on the spectrograms represents formants that match a vocal tract resonance, and the formants were measured using linear frequency coding (LPC) tracks as done in other studies (e.g., Khulage & Pathak, 2012).

2.9 Studies on the acoustic analysis of vowels

Studies on the acoustic analysis of vowels commonly measure the first two formants of the vowels, although there are some which measure three (e.g., Gold & Earnshaw, 2019). There are two main ways to measure the vowels. One is two measure the vowels at intervals (e.g., Gold and Earshaw, 2019), and the other is to measure the vowel at its midpoint (e.g., Duniec & Crouzet, 2014).

The values of the formants in Hertz (Hz) are generally converted to an auditory Bark scale. The purpose of converting these values is so that "the distance between formant values on the plot might be similar to the way that distances in vowel quality are actually perceived" (Deterding, 2003, p. 4). The formula to convert the Hz values to the Bark scale was suggested by Zwicker and Terhardt (1980, p. 1524).

$$\text{Bark} = 13 \arctan(0.76\, f / 1000) + 3.5 \arctan(f / 7500)2.$$

2.10 Conclusion

This chapter discussed the classification of Austronesian and Austroasiatic languages, language endangerment and OA languages as well as the Malay borrowings in OA languages. In order to compare the vowel inventories between other OA languages and Kensiu, the previous studies of OA sounds were discussed. The explanation on Kensiu sounds spoken in Yala Province were described in order to compare them with the Kensiu variety in Kedah, Kedah Malay and also standard Malay. Lastly, this chapter also discussed the formant measurements of the vowels and the studies on the acoustic analysis of vowels.

CHAPTER 3
METHODOLOGY

3.1 Introduction

This chapter provides an explanation of the methods used in this study. The first part covers the selection of speakers, followed by the instruments and materials used in this study. Following this, an explanation of the data collection procedure, and the analysis of the data are provided. The last section in this chapter defines the key terms used in this study.

3.2 Selection of the speakers

The sampling technique used in this study was a non-probability sampling which Showkat and Parveen (2017) define as a technique that involves non-randomized methods to select the sample. In non-probability sampling, the selection of the speakers is based on the availability and the convenience of the speakers (Creswell, 2008). This study used a snowball and convenience sampling approach to obtain participants. This approach was chosen as this study focused on a group of minority people who speaks an endangered language. In this case, a random selection method was not suitable as nonfluent speakers may be selected. As mentioned in Chapter 1 (see 1.3) Kensiu speakers tend to use more Kedah Malay than Kensiu, especially among the younger generation. Snowball sampling

was deemed to be the best method to use given that the researcher is not a part of the community. Thus, having an intermediary helped in the selection of participants that fulfil the desired criteria of the study. The selection of the speakers in this study was helped by P1, the first participant, during the second visit at the site. The request for help was done through informal conversations with the participants

(Creswell, 2008). In order to ensure that the selected participants fulfilled the requirements set, mock recordings were carried out. The mock recordings were carried with the potential participants by explaining to them the data collection procedures.

In terms of the sample size, only five participants from Kampung Siong, Lubok Lengong (° 47' 57" North, 100° 54' 51" East), Baling Kedah were recruited in this study. Kampung Siong was chosen for the study because the concentration of Kensiu people is high there in spite of the fact that several numbers of Kensiu can also be found in Perak (see 1.3). Although the sample size was small, it represents 2.5% of the Kensiu population. The intention of this study was not to generalize the data to the entire population but to provide an understanding of the central phenomenon.

Convenience sampling based on a set of criteria was used to determine the suitability of participants for this study. This was to ensure the homogeneity of the participants. The first criterion to participate in this study was that they had to be females. Due to biological differences in the vocal tract between males and females, only one gender was chosen in this study. As indicated by Simpson (2009), male and female speech differs in their phonation and pitch, articulation, articulatory dimensions, and articulatory speech. In order to make sure that the measurements of the frequencies were consistent and did not affect the measurements of the vowels, the gender variable was kept constant. Besides the uniformity of their biological traits, it was noticed that it was easier to build a close relationship between the researcher and the participants as they were more comfortable to share their personal experiences and thoughts with a person who was of

the same gender as them. Female participants were also chosen due to the consideration that they were available in the village and has less influence from the outside world compared to the men who went out to work.

The second criterion was age. All the chosen participants were aged between 50 to 62 years old with an average age of 56.4 years. By selecting this age group, the data were more reliable in comparison to recordings of speech from the younger generation who had a stronger tendency to code-mix Kensiu with Kedah Malay. The reason for this is that they use Malay to interact with the majority Malay school population, and because they live in Kedah, they are more likely to use the local Kedah Malay dialect. Hence, the remaining speakers of Kensiu who speak the language fluently tended to consist of older speakers aged 50 years old and above. During the first visit to the village, a few of the Kensiu residents also said that they often needed to switch to Kedah Malay when speaking with their children as some of them could not understand some of the words and phrases in Kensiu. This suggests that Kedah Malay may emerging as the first or dominant language among younger Kensius.

The third criterion was related to the second criterion that was taken into account, which is that all the participants must speak Kensiu as their first language and use mostly Kensiu in their daily communication. Ladefoged (2003) stated that the first step when selecting the participants is to make sure that the language they use is their mother tongue. It was observed that in the Kensiu settlement, the women often had frequent contact with the Malay community since their settlement was located near to a Malay village and they often buy their daily products at the local Malay market. Thus, it can be said that they also used Kedah Malay regularly through the interaction with the Malay community. However, they used Kensiu predominantly amongst themselves and declared it to be their mother tongue.

The last criterion was that the speakers had no speech and hearing problems. Since this study only employed participants from a senior age group, the physical condition of the participants was taken into consideration. The selection was carried out carefully as some of the older speakers had physical disabilities, such as teeth loss and poor hearing, and had illnesses that might hinder them from producing clear speech. All the speakers selected for this study had no speech or hearing impediments.

3.3 Consent

Necessary permissions were obtained prior to collecting the data. The approval to carry out this study was first obtained from the Department of Orang Asli

Development (JAKOA) in May 2019 before the first visit in June that year (see Appendix 2). Informed consent to record the speakers and to use the recordings for this study and other related presentations and publications was obtained from each participant after explaining to them the nature of the study and informing them of their right to withdraw from the study (see Appendix 1). Since the speakers' level of literacy differed, permission was sought orally from the participants after reading the contents of the consent form, which was in Malay, to them. The participants acknowledged that they understood the objectives of the study and expressed their willingness to take part in the study (Creswell, 2008). Verbal permission to use the community hall in the village was sought from Encik Razali Bin Kulim, who is the village head or *Tok Batin*.

3.4 Materials and instruments

The following sections explain the materials and instruments used in this study.

3.4. 1 Demographics interview

Demographic information was gathered from each participant by recording them orally in Kedah Malay. To obtain demographic information, the participants were asked about their basic information, such as age, religion, level of education, occupation, and languages spoken in the village and outside the village (see Appendix 3 for the questions asked to elicit demographic information). Specific information, such as the level of education, was included because their language might be different if their education level was higher. Some of the participants attended primary school, and hence, they can read and write in Malay (basic level). All the participants have not and do not work outside their village, meaning that in their living environment, they would have less contact with the outsiders who speak other languages. However, they often purchased fresh produce at the local Malay market. That explains the language spoken by the Kensius to people inside the village and outside the village, which is Kedah Malay. Since almost half of the villagers are Muslims, some of the participants spent their time in Quran recital classes once a week, which is carried out in Quranic Arabic but the religious leader, who is Malay, uses Malay during class interaction.

In this study, the participants are coded as P1 until P5. Their demographic information is shown in Table 3.1.

Table 3.1. Demographic background of Kensiu speakers

Speakers	Age	Religion	Spoken Language		Occupation	Level of education
			Inside the village	Outside the village		
P1	60	Islam	Kensiu	Kedah Malay, Baling dialect	Housewife	Primary
P2	59	Islam	Kensiu	Kedah Malay, Baling dialect	Housewife	-
P3	50	Islam	Kensiu	Kedah Malay, Baling dialect	Housewife	Primary
P4	51	Islam	Kensiu	Kedah Malay, Baling dialect	Housewife	-

P5	62	Islam	Kensiu	Kedah Baling dialect	Malay,	Housewife -

3.4.2 Selection of words and pictures

The selection of words containing front oral monophthongs was extracted from the glossary provided by Bishop (1996). The reason to adapt from another Kensiu variety is that there are no published word lists on the variety spoken in Kedah related to the production of vowels. The choice of the lexical items from Bishop's (1996) glossary was based on the vowel being in a closed syllable due to the reason that an open syllable might affect vowel lengthening (Rawlins, 2006). There are three types of syllables in the word list, such as CVC *tek* [tẹk] 'soil', CVCVC *hetit* [hətit] 'tail', and VCVC *ages*

[ages] 'mosquitoes'.

As far as possible, the vowels were preceded or followed by plosives and fricatives while nasals, liquids, and approximants were avoided. The rationale for avoiding these was to minimize co-articulatory influences of the subsequent word (Tunley, 1999). However, there were some cases where this could not be avoided, where the target vowels were followed or preceded by nasals and approximants, such as the words in *gunting* [guntɪŋ] 'scissors', *hewit* [həwɪt] 'throw-away', *gading* [gadɪŋ] 'trunk' and *ages* [ages] 'mosquitoes'. After the lexical items were selected from

Bishop's (1996) glossary, these items were sorted based on their word classes. Table 3.2 provides the list of words used in this study as well as their word class.

Table 3.2. List of target words

Target vowel	Transcription		Malay	English	Word Class
	Bishop (1996)	Speaker's pronunciation			

/i/	/hətit/	/hetik/	ekor	tail	noun
	/lətik/	/lətik/	lidah	tongue	noun
	/dakiʔ/	/dakik/	daki	dirt	noun
	/kəsiʔ/	/kəsik/	sunyi	lonely	adjective
	/nasiʔ/	/nasik/	nasi	rice	noun
	/bəhiʔ/	/bəhik/	kenyang	full	verb
	/tis/	/tis/	cendawan	mushroom	noun
	/ʔis/	/is/	lintah	leech	noun
/ɪ/	/guntɪŋ/	/guntiŋ/	gunting	scissors	noun
	/həwɪt/	/həwit/	buang	throw	verb
	/jəlabɪt/	/dʒəlabit/	melekit	sticky	adjective
	/gadɪŋ/	/ gadɪŋ/	gading	trunk	noun
/ẹ/	/ʔẹk/	/ek/	anjing	dog	noun
	/tẹk/	/teʔ/	tanah	soil	noun
	/yakẹt/	/yakit/	rakit	raft	noun
/e/	/hubet/	/hubet/	ubat	medicine	noun
	/tənapes/	/tənapes/	penapis	sieve	noun
	/ages/	/ages/	nyamuk	mosquitos	noun
/ɛ/	/gɛhit/	/gɛhet/	pahit	bitter	adjective
	/klapɛh/	/klapɛh/	lengan	arm	noun
	/ləbɛh/	/ləbeh/	buluh	bamboo	noun
	/bəkɛs/	/bəkes/	bekas	food container	noun
	/bəjakɛs/	/bəjakes/	dewasa	adult	noun
	/gɛs/	/ges/	dapur gas	gas	noun

Target vowel	Transcription		Malay	English	Word Class
	Bishop (1996)	Speaker's pronunciation			

According to Simons and Fennig (2017), Kensiu uses Latin and Thai script as its writing system. However, Duangchan (2006) raised the issue that orthographies are not widely used in the languages of northern Aslian in Southern Thailand including the variety which is spoken in Thanto District, Yala Province. Since Kensiu is a spoken language with a limited writing system, all the words chosen were presented in pictures as a better way of eliciting data. Not only does Kensiu not have its own set of conventions for its language, but not all the participants were literate. Thus, pictures were thought to be the best way to use with participants to elicit the data. There were twenty-four sets of pictures presented to the participants and each picture depicted a word containing the target vowel (see Appendix 5 for samples of the pictures used).

3.4.3 Instruments

The speakers were recorded using a *Zoom H6 Handy Recorder* and head worn *Audio Technica* microphone. The sampling rate was set as 44.1 kHz with 16-bit resolution. The recording session was held in the community hall in the village. Ambient noise was controlled by recording the speakers in the hall as the windows and doors could be closed. The recordings were also done in the middle of the hall away from the windows. The important thing to be careful about when doing the recording in the hall was the background noise and echo, and the best way to reduce these is by putting the microphone in close proximity approximately 5cm from the participants' lips to the side (Butcher, 2013). As suggested by Butcher (2013), a head-mounted microphone is the best instrument to be used when the noise in the indoor recording is unavoidable. This is

supported by Kent and Read (2002) who advise that using a headmounted condenser microphone is a good instrument as it is still can maintain the quality of the recordings even when the speakers move their body and head position. Although the recordings were done in a hall, it was still impossible to make the recordings completely noise-free as there were always vehicles passing by the main road by the village and children sometimes played outside the hall. Figure 3.1 shows one of the participants testing the instruments before the actual recordings started.

Figure 3.1. Testing the instruments with a participant (photograph used with permission)

3.5 Data collection procedure

Phonetic fieldwork requires a great amount of patience and determination as it can be very challenging, exhausting and time-consuming. Thorough planning before the trip is crucial to ensure that the researcher can establish a good relationship with the community. However, as put forward by Gordon (2003) detailed preparation by the researcher per se may not guarantee that the work would turn out as expected. This is supported by Butcher (2013) who maintains that fieldwork also depends on good luck rather than a great preparation. During the initial visit to Kampung Siong, the researcher

was warmly welcomed by the *Tok Batin* and his family. During the conversation, he shared the general overview of Kensiu people specifically their background of the economic status and their lifestyle. He also raised his concerns and worries that the Kensiu language is getting less spoken among their youth nowadays. Getting to know the community members was a challenging task. However, the *Tok Batin* was kind enough to assist the researcher during the process of selecting the participants. On a second visit, few members were approached and through the first introduction, it was easier to expand the network. The researcher casually interviewed the participants about their usual lifestyle to make sure their daily activities were not interrupted once the recordings start. The schedule for the recording session was planned with the agreement from the researcher and the participants. Respecting their time is one of the ways to show respect, and minimises intrusion by the researcher. In this study, the recordings were done twice a week for around two hours for each session.

At the beginning of the recordings, before the final word list was finalized, the researcher went through all the words in the Bishop's (1996) glossary with the speakers as the list use is based on Kensiu in Yala Thailand. It was necessary to get the list checked because the lexical items might be different between the two varieties. The data were obtained by presenting the pictures in a laptop computer instead of the printed version of pictures as the size of the pictures and the level of brightness could be adjusted. The pictures were arranged according to their word classes to avoid confusion among the participants. Coupe (2014) points out that misinterpretations and confusion can easily occur if the list of words jumps from one-word category to another. Thus, arranging the lexical items based on their word-class allowed the participants to recognize the recurring morphology (Coupe, 2014). Most of the words obtained were nouns, adjectives, and the least were verb (see Table 3.2). This is due to a reason that, the researcher had to give additional explanations on the verbs and adjectives because it may not be reflected in the

pictures. On the other hand, nouns were easier as the speakers could straight away provide the description for the pictures by looking at the objects shown. The speakers were audio-recorded describing the pictures. For example, to elicit the word /ages/ 'mosquitoes', the researcher asked the speaker "*Makcik, nyamuk dalam bahasa Kensiu panggil apa ya?"* while showing a picture of mosquitoes.

When the speaker uttered the word 'ages', she was then asked to construct a sentence using the word in Kensiu. The speakers were encouraged to construct the sentences as naturally as possible, that is, the way they use Kensiu in their daily interactions. This was done so that the speech would sound more natural as they would be less aware in attempting to construct the sentences carefully.

The description of the pictures also provided a context for the word meaning that they were not produced in isolation because it may affect the word stress and intonation. They were asked to repeat the description twice. This way, there were two samples for each recording. Gordon (2003) stated that every word should at least be recorded twice to avoid any background noise that might happen or speech dysfluencies from the speakers. It was also to make sure that the first pronunciation is not a coincidence. It was important to ask the speakers to translate their descriptions as well after each recording, so the researcher can understand the meaning of the whole sentence rather than the meaning of the target word only.

3.6 Data analysis procedure

A total of 203 tokens were used in the analysis. Some tokens were eliminated due to the noise in the recordings. As mentioned earlier (see 3.4.2), it was not always possible to avoid nasals and approximants neighbouring the vowels. This was especially for the vowel /ɪ/. Thus, the measurements had to be more carefully carried out to ensure that they

were not affected by the neighbouring nasals and approximants. Table 3.3 shows the number of tokens recorded from five participants.

Table 3.3. Number of tokens in each participant

Target vowel	Number of tokens elicited in each participant					TOTAL
	P1	**P2**	**P3**	**P4**	**P5**	
/i/ 15	11	10	13	14	63	
/ɪ/ 7	8	6	7	6	34	
/ẹ/ 6	6	5	5	4	26	
/e/ 6 6 6 6 6 30						
/ɛ/ 11 9 8 10 12 50						

The translations of the sentences were done by each participant, so the translations would be more accurate. The representation of Kensiu words in this study was presented in a Malay-based phonemic spelling system where the words were spelled close to how they were produced as is done in the Malay spelling system. The same method also used by Baxter and de Silva (2004) for Malacca Portuguese. A phonemic system was used because the orthography system of some OA languages still lacks accuracy and practicality as they are not sufficiently analysed (see Benjamin, 2012a). The list of words, transcriptions, and translations can be seen in APPENDIX 4.

The recordings obtained were analysed using PRAAT 6.1.08 (Boersma & Weenink, 2019). The sound files were then annotated in TextGrid in PRAAT, the orthographic transcriptions of the target words were located on the first tier and the target vowels were located at the second tier. After the vowels were segmented, the F1 and F2 were measured at the centre of the vowels as the vowels were at the stable state

(Derdemezis et al., 2016). The F1 and F2 were generated automatically using PRAAT scripts. Praat script is a plain text file that has a series of commands to PRAAT (Styler, 2017). However, manual measurements were needed when the vowels seemed odd or where there were, for example, nasal consonants in neighbouring environments. The generated measurements from PRAAT scripts were then transferred into an excel file to calculate their values of F1 and F1 averages, and F1 and F2 standard deviations.

The measurement of each vowel in Hz and Bark values are provided in APPENDIX 6. The rationale for measuring the F1 and F2 is based on the formant frequency model (see 2.8 and 2.9), where F1 is inversely related to vowel height and F2 in related to vowel advancement/ retraction (Kent and Read, 2002). The values of the formants in Hz entered were then converted into Bark values using the formula given by Zwicker and Terhardt (1980, p. 1524) (as discussed in 2.9).

$$\text{Bark} = 13 \arctan (0.76 f / 1000) + 3.5 \arctan (f / 7500)2.$$

The formant values were converted into a Bark scale as the distance of the formant values on the plot is similar to the distance of vowel quality in the vowel space (Deterding, 2003). The values of F1 were plotted on the y-axis and values of F2 were plotted in the x-axis. Pillai, Zuraidah Mohd Don, Knowles, and Tang, (2010, p. 164) stated that "the F1 vs. F2 plot, in any case, gives a better representation of the traditional vowel quadrilateral than F2 vs. F1 plot". Apart from determining F1 and F2 to identify the characteristics of the vowels (see research question 1), ANOVA was carried out to find the difference between the speakers. Lastly, two-tailed independent *t-test* was carried out to find the significant difference between the average values of each vowel to examine if the vowels were contrasted (see research question 2). Previous studies such as Hillenbrand and Clark (2000) and Clopper, Pisoni and Jong (2005) have used these tests to measure the acoustic characteristics of the vowels. However, the results of this present study should be treated with caution as the result could not be generalized to the whole Kensiu's population. Thus, the results may be valid only for five participants recruited in this study.

3.7 Operational definition of terms

This dissertation consists of a few terms used to measure and analysing the data that may be unfamiliar to certain readers. The following terms are operationally defined based on the usage in this study.

a) Significant and not significant

In this study, the term significant is used to discuss the results from statistical analysis. The result is significant when the p-value is less than 0.50 and the result is not significant when the p-value is equal or more than 0.50.

b) Small, medium and large effect size

The term small, medium and large effect sizes were used to explain the results for effect size based on Cohen's value. Small effect size means the value *d* is less or equal than 0.20, $d \leq 0.20$. Medium effect size refers to the *d* value which is equal to 0.50, $d = 0.50$. Large effect size refers to the *d* value which is greater or equal than 0.80, $d \geq 0.80$.

e) Vowel conflation

Vowel conflation in this study refers to the merging of two vowels when they possess the same acoustic properties.

3.8 Conclusion

This chapter discussed the methodology used in this study. The next chapter provides and discusses the findings of the study.

CHAPTER 4
FINDINGS AND DISCUSSION

4.1 Introduction

In this chapter, results from the F1 and F2 measurements of the Kensiu front oral vowels will be presented in several parts. The first part will look at the overall placement of these vowels. This will be followed by the findings pertaining to each vowel and comparisons of similarly produced vowels. The results will then be discussed and compared to the description of Kensiu vowels by Bishop (1996), Kedah Malay and Standard Malay.

4.2 Vowel quality of front oral vowels in Kensiu

The total number of tokens for /i/, /ɪ/, /ẹ/, /e/ and /ɛ/ were 63, 34, 26, 30 and 50 respectively. Table 4.1 shows the average values and the standard deviation (SD) for F1 and F2 in Hertz and the F1 and F2 values in Bark. The measurement for each vowel is provided in APPENDIX 6.

Table 4.1. Average values of F1 and F2 and standard deviation

Vowel	Ave. F1 and SD (Hz)	Ave. F2 and SD (Hz)	Ave. F1 (Bark)	Ave.F2 (Bark)

/i/	358.56	2637.95	3.46	14.84
	(39.71)	(237.81)		
/ɪ/	358.74	2582.68	3.47	14.71
	(40.62)	(322.12)		
/ẹ/	422.00	2362.10	4.05	14.16
	(43.74)	(196.89)		
/e/	491.90	2293.10	4.67	13.97
	(89.58)	(257.47)		
/ɛ/	545.16	2262.36	5.12	13.89
	(103.72)	(299.35)		

SD = Standard Deviation Ave = Average

From Table 4.1 it can be seen that the highest average of F1 value was for /ɛ/ while lowest F1 value was /i/. This is consistent with the height of these two vowels in the vowel space with /i/ being a higher vowel than /ɛ/. In terms of the second formant, the highest average F2 value was for /i/ and the lowest one was /ɛ/. However, the differences were not big given that all the vowels in this study were front vowels.

As mentioned in Chapter 2, Bishop (1996) describes Kensiu as having five front oral vowels. However, as can be seen in Figure 4.1, /i/ and /ɪ/ appear to be overlapping. However, on the whole, the similarities of the vowel positions for each vowel between the variety of Kensiu in Yala and the one in Kedah are obvious. Bishop (1996, p. 229) described /i/ as being placed "slightly higher than its corresponding /ɪ/ which in turn is described as being "slightly fronted". At a glance, this may be the case as shown in

Figure 4.1. Second, consistent with Bishop (1996, p. 229), /e/ was located "lower and more backed than /ẹ/" which is described as being a "close mid-front" vowel while /e/ is described as a "mid-front" vowel. However, as can be seen in Figure 4.1, the position of /e/ was slightly more retracted than /ẹ/. As would be expected, the vowel /ɛ/ was placed lower than the other vowels. Bishop (1996, p. 229) described it as an "open-mid front" vowel although the vowel in Figure 4.1 appears to have been produced higher. Figure 4.1

illustrates the location of each of the front vowels in Kensiu in the vowel space. This figure also reflects the expected positions of the front vowels in the vowel space.

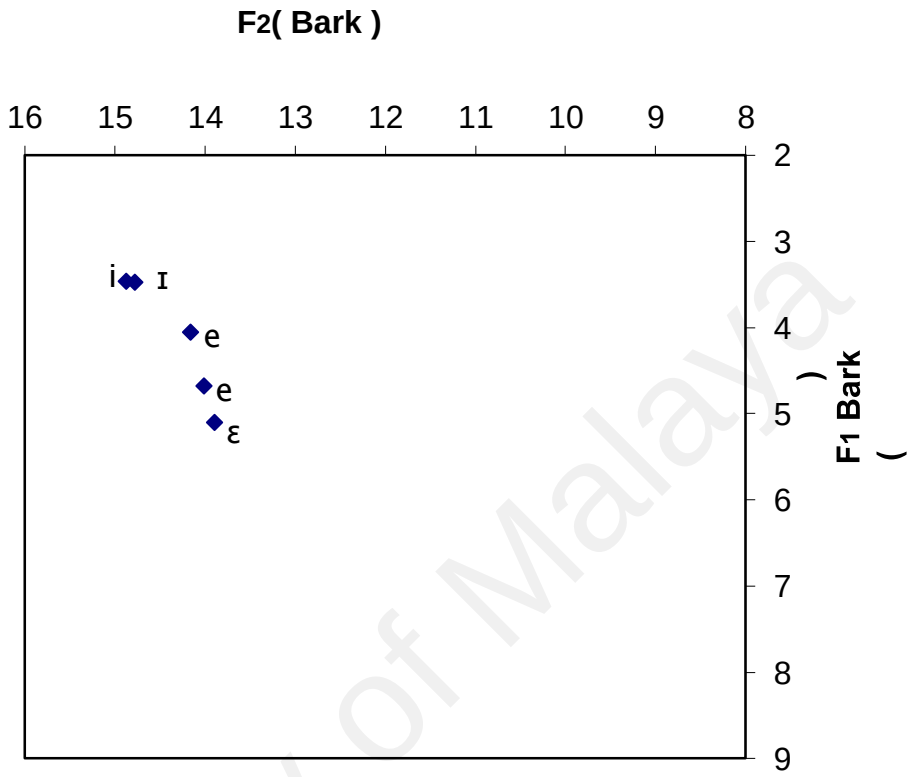


Figure 4.1. Vowel plot of Kensiu front vowels

4.2.1 Kensiu /i/

The vowel /i/ was extracted from eight words which were *hetit* [hətit] 'tail', *letik* [lətik] 'tongue', *daki* [dakiʔ] 'dirty', *kesi* [kəsiʔ] 'lonely', *nasi* [nasiʔ] 'rice', *behi* [bəhiʔ] 'full', *tis* [tis] 'mushroom' and *is* [ʔis] 'leech'(see Table 3.2). It should be noted that Kensiu words often have a glottal stop'ʔ' in word-final and word-initial positions before and after a vowel, for example, [hetiʔ] and [ʔis]. Table 4.2 shows the frequency of /i/ produced in each word.

Table 4.2 Frequencies of /i/ in each word

Kensiu word	Meaning	Frequency

[nasiʔ]	rice	9
[hətit]	tail	10
[lətik]	tongue	10

Kensiu word	Meaning	Frequency
[dakiʔ]	dirty	4
[kəsiʔ]	lonely	4
[bəhiʔ]	full	7
[tis]	mushroom	9
[ʔis]	leech	10

Figure 4.2 illustrates the distribution of /i/ produced by the five Kensiu speakers. The dispersion of /i/ has the lowest standard deviation which is 113.80 Hz compared to other vowels. Hence, it can be surmised that they were produced in a similar manner and this is shown in the clustering of all the tokens of /i/ produced by the speakers. As can be seen in Figure 4.2, the realisations of these vowels were generally clustered together at the close-front position. The small variation of the production of /i/ suggests that this is a stable form of Kensiu /i/ where each participant produced the same sounds with a small level of variation. Speaker P5 produced a more fronted /i/ that the rest of the productions. Speaker P2's production of /i/ appeared to be lower than the others and the speaker presented in red (P4) seemed to have produced the vowel higher. Speaker P1, represented in blue, produced a more retracted /i/. The /i/ produced by speaker P3 were mostly scattered in the middle of the vowel space. However, there was an instance for P3 that was located away from the cluster. The word uttered was *tis* [tis]

'mushroom' in the first elicitation. Figure 4.2 displays the distribution of /i/ in the vowel space.

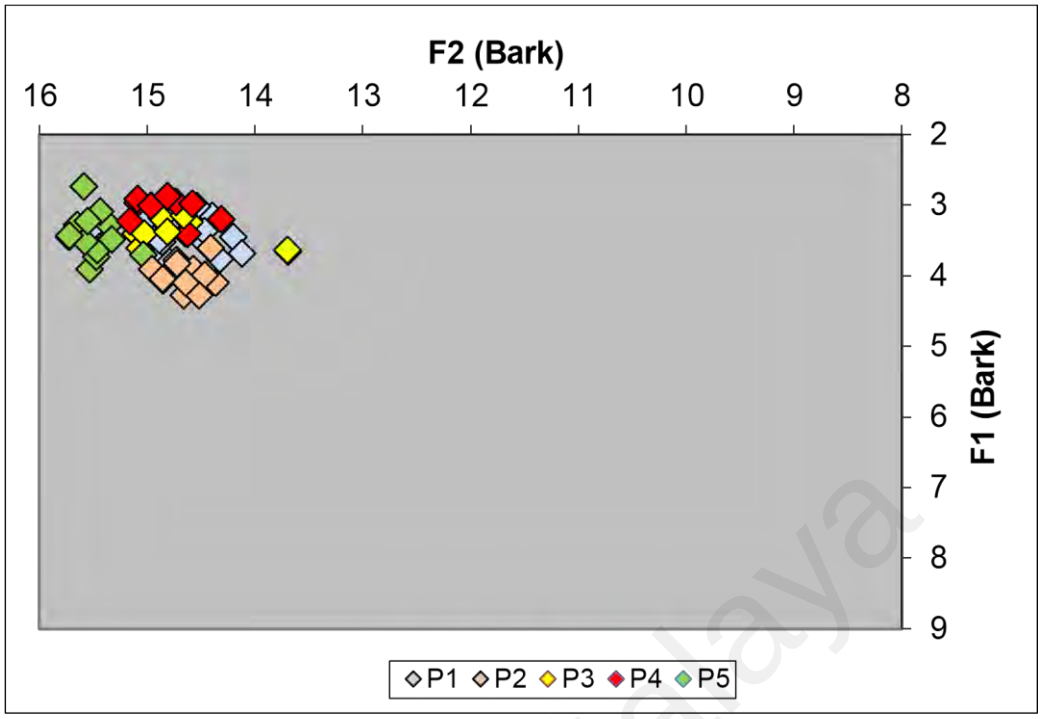


Figure 4.2. Scatter plot of /i/ in Kensiu

A one-way ANOVA was carried out to test the differences among the five speakers. The results indicated that there were significant differences between the average F1 values of the five speakers, $F(4, 58) = 22.06$, $p < 0.05$. A Tukey's HSD showed that F1 means were significantly different between all speakers except for P1

($M = 350.47$ Hz, $SD = 29.73$) and P3 ($M = 356.00$ Hz, SD = 21.06), P1 ($M = 350.47$ Hz, $SD = 29.73$) and P5 ($M = 355.14$ Hz, $SD = 32.17$), and P3 ($M = 356.00$ Hz, $SD = 21.06$) and P5 ($M = 355.14$ Hz, $SD = 32.17$).

A one-way ANOVA showed that there were significant differences between the average F2 of the five speakers, $F(4, 58) = 15.16$, $p < 0.05$. A Tukey's HSD test revealed that F2 means were significantly different between the groups except for P1 (M

= 2543.53 Hz, $SD = 108.97$) and P2 ($M = 2540.27$ Hz, $SD = 76.71$), P1 ($M = 2543.53$ Hz, $SD = 108.97$) and P3 ($M = 2480.70$ Hz, $SD = 371.53$), P1 ($M = 2543.53$ Hz, $SD = 108.97$) and P4 ($M = 2622.00$ Hz, $SD = 109.07$), P2 ($M = 2540.27$ Hz, $SD = 76.71$) and

P3 (*M* = 2480.70 Hz, *SD* = 371.53), P2 (*M* = 2540.27 Hz, *SD* = 76.71) and P4 (*M* = 2622.00 Hz, *SD* = 109.07), and P3 (*M* = 2480.70 Hz, *SD* = 371.53) and P4 (*M* = 2622.00 Hz, *SD* = 109.07). Based on the F2 measurements, the productions of /i/ were more dispersed in the vowel space. Table 4.3 shows the formant measurements of /i/ for each speaker.

Table 4.3. Formant measurements for Kensiu /i/ in each speaker

Speaker	Ave. F1 and SD (Hz)	Ave. F2 and SD (Hz)
P1	350.47	2543.53
	(29.73)	(108.97)
P2	419.27	2540.27
	(21.55)	(76.71)
P3	356.00	2480.70
	(21.06)	(371.53)
P4	322.15	2622.00
	(18.97)	(109.07)
P5	355.14	2943.00
	(32.17)	(89.50)

4.2.2 Kensiu /ɪ/

The vowel /ɪ/ was extracted from words, such as *gunting* [guntɪŋ] 'scissors', *hewit* [həwɪt] 'throw-away', *jelabit* [jəlabɪt] 'sticky', and *gading* [gadɪŋ] 'trunk'(see

Table 3.2). There were three occurrences of /ɪ/ between approximant and nasals which were *gunting* [guntɪŋ] 'scissors', *gading* [gadɪŋ] 'trunk', and *hewit* [həwɪt] 'throwaway'. As mentioned in Chapter 3 (see 3.5), these vowels were re-measured using visual and auditory inspection of the spectrograms. Table 4.4 shows the occurrences /ɪ/ in each word.

Table 4.4. Frequencies of /ɪ/ in each word

Kensiu word Meaning Frequency [guntɪŋ] scissors 9		
[həwɪt]	throw-away	7
[jəlabɪt]	sticky	8
[gadɪŋ]	trunk	10

Figure 4.3 displays the distribution of the production of /ɪ/ by Kensiu speakers. The distribution of /ɪ/ was scattered between a close and front position. In terms of vowel fronting, these vowels were highly dispersed compared to /i/, but in terms of vowel height, these vowels appear to be more concentrated at the close position like /i/. As shown in Figure 4.3, again, similar to the vowel distribution of /i/, speaker P5 produced a more fronted vowel than the rest of instances. Vowels elicitations by P1 were more concentrated in the middle. It also can be seen that P3's production of /ɪ/ was slightly dispersed to the back. Meanwhile, speaker P4 seemed to produce higher vowels compared to P2 who produced this vowel lower than the others. However, there were two instances of /ɪ/ by P4 that were located away from the rest of her realisations which clustered at the close and front position. These two instances were *hewit* [həwɪt] 'throwaway' for the first and second elicitation. The measurements were checked and were noted as being correct. It was noticed that when /ɪ/ was pronounced after the approximant /w/, the vowel seemed to be more retracted perhaps due to the influence of the velar approximant. Hence, this could explain the two realisations of /ɪ/ by speaker P4 which were located further back than the other productions. Figure 4.3 illustrates the distribution of /ɪ/ in the vowel space.

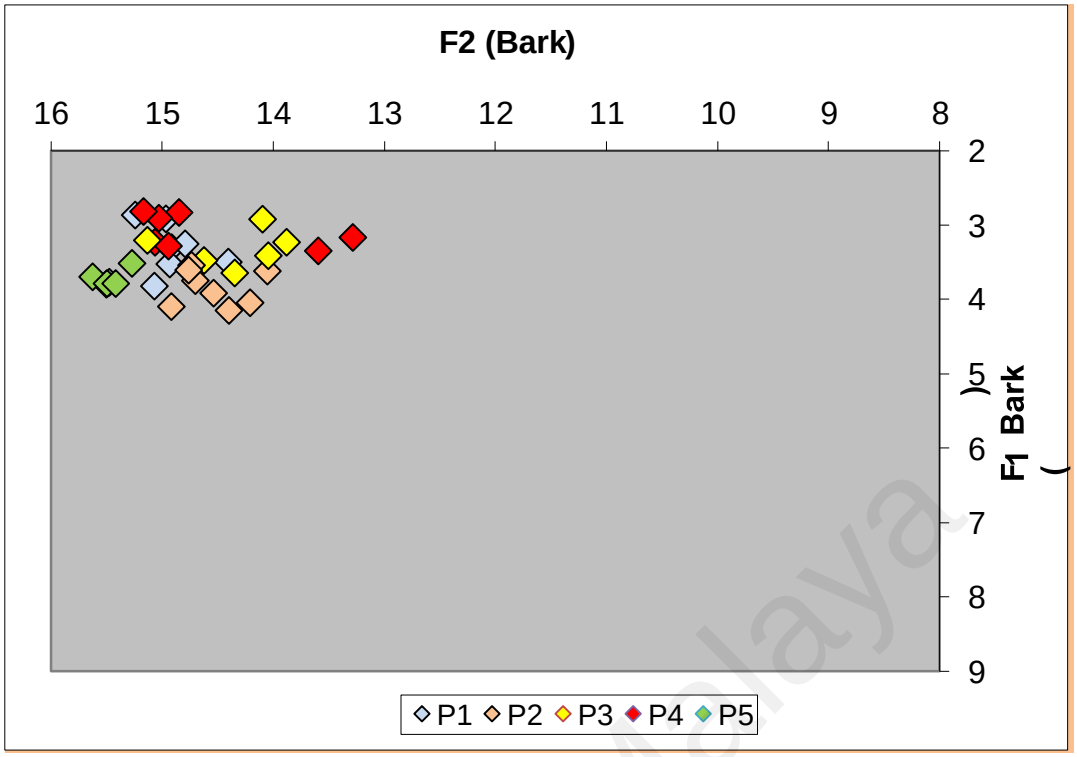


Figure 4.3. Scatter plot of /ɪ/ in Kensiu

A one-way ANOVA showed that there were significant differences between the F1 average values of the five speakers, $F(4, 29) = 11.57$, $p < 0.05$. A Tukey's HSD test showed that the F1 means were not significantly different among speakers P1 ($M = 341.43$ Hz, $SD = 36.68$) and P3 ($M = 343.17$ Hz, $SD = 27.18$), P1 ($M = 341.43$ Hz, $SD =$ 36.68) and P4 ($M = 318.14$ Hz, $SD = 24.13$), P2 ($M = 400.13$ Hz, $SD = 26.81$) and P5 ($M = 386.67$ Hz, $SD = 11.72$), P3 ($M = 343.17$ Hz, $SD = 27.18$) and P4 ($M = 318.14$ Hz, $SD = 24.13$), and P3 ($M = 343.17$ Hz, $SD = 27.18$) and P5 ($M = 386.67$ Hz, $SD = 11.72$). Thus, it can be concluded that most of the vowels were produced similarly in terms of the vowel height.

A one-way ANOVA showed that there were significant differences between F2 average values of the five speakers, $F(4, 29) = 5.08$, $p < 0.05$. Tukey's HSD test revealed that there were no significance differences between speakers except between the speakers P2 ($M = 2671.29$ Hz, $SD = 113.10$) and P5 ($M = 2930.83$ Hz, $SD = 57.24$), and P3 ($M =$

2272.33 Hz, *SD* = 504.64) and P5 (*M* = 2930.83 Hz, *SD* = 57.24). As for the vowel fronting between the groups, only P2-P5 and P3-P5 produced the vowels differently. Table 4.5 shows the average formant measurements for /ɪ/ in each speaker.

Table 4.5. Formant measurements for Kensiu /ɪ/ in each speaker

Speaker	Ave. F1 and SD (Hz)	Ave. F2 and SD (Hz)
P1	341.43	2671.29
	(36.68)	(113.10)
P2	400.13	2513.88
	(26.81)	(121.18)
P3	343.17	2272.33
	(27.18)	(504.64)
P4	318.14	2540.29
	(24.13)	(301.23)
P5	386.67	2930.83
	(11.72)	(57.24)

4.2.3 Kensiu /ẹ/

The vowel /ẹ/ was taken from three words which are *ek* [ʔẹk] 'dog', *tek* [tẹk] 'soil', and *yaket* [yakẹt] 'raft' (see Table 3.2). As mentioned earlier in this chapter, the glottal stop usually appears at the beginning or the end of the word. Thus, *ek* [ʔẹk] 'dog' was considered as closed syllable as there is a glottal stop preceding /ẹ/. Table 4.6 shows the frequencies of /ẹ/ in each word.

Table 4.6. Frequencies of /ẹ/ in each word

Kensiu word	Meaning	Frequency
[ʔẹk]	dog	7
[tẹk]	soil	10

Figure 4.4 shows the distribution of /ẹ/ based on the production of this vowel by each participant. This figure shows that there is a higher level of overlap at the midclose and mid-front positions. Speaker P2 produced lower vowels than the other instances except for the second elicitation in word *yaket* [yakẹt] 'raft' that was produced higher than the rest of P2. Meanwhile, the most consistent production of /ẹ/ was by P1 where the vowels were all clustered together in the middle of the vowel space. Speaker P5's production tends to be more fronted, however, there were not enough tokens collected for P5, and thus there is no strong evidence in terms of the way they arrived.

Speaker P3's realisations of this vowel were away from the area in the vowel chart, that is, where the realisations by the other speakers were clustered. There were five instances which speaker P3 was trying to emulate: *tek* [tẹk] 'soil' in first and second elicitation, *ek* [ʔẹk] 'dog' for the first elicitation and *yaket* [yakẹt] 'raft' for the first and second elicitation. As can be seen in Figure 4.4, the first elicitation of *yaket* by P3 was positioned at the close and front position in the chart. The recording of *yaket* [yakẹt] 'raft' first elicitation was checked and it was found that the P3 speaker pronounced *yaket* as *yakit* in /i/ vowel rather than /ẹ/. Thus, that is why it was located in a high front position like /i/. In order to confirm P3's production, where all the vowels more retracted than the rest, these vowels were checked and, but their measurements remained the same. This could be due to the speaker's pronunciation of the word which was different from the ways in which the other pronounced it. Figure 4.4 shows the distribution of /ẹ/ in the vowel space.

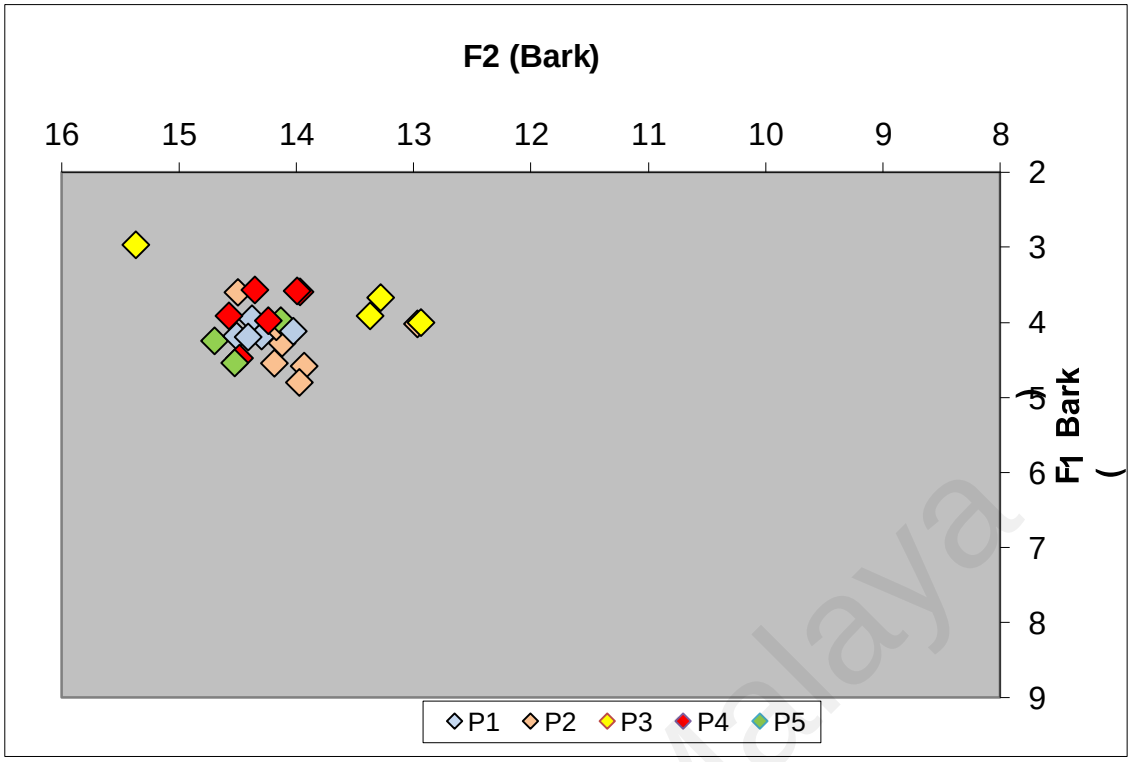


Figure 4.4. Scatter plot of /ẹ/ in Kensiu

A one-way ANOVA of four speakers was performed, and significant differences between the F1 average values were found: $F(3, 18) = 4.71$, $p < 0.05$. P5 was removed from the analysis as there were less than five instances from the data. Tukey's HSD test results show that the F1 means were significant differences only between P2 (M = 454.83 Hz, SD = 47.22) and P3 (M = 386.6 Hz, SD = 47.46), and P2 (M = 454.83 Hz, SD = 47.22) and P4 (M = 387.8 Hz, SD = 22.30). This indicates that the three speakers produced the vowel /ẹ/ differently in terms of vowel height.

A one-way ANOVA showed that there were no significant differences between their F2 means, $F(3, 18) = 1.29$, $p = 0.31$. Tukey's HSD results also showed that there were no significant differences between any two groups meaning that all four speakers produced /ẹ/ similarly in terms of vowel height. Table 4.7 shows the formant values for /ẹ/ vowel for each speaker.

Table 4.7. Formant measurements for Kensiu /ę/ in each speaker

Speaker	Ave. F1 and SD (Hz)	Ave. F2 and SD (Hz)
P1	425.67	2393.5
	(11.54)	(70.52)
P2	454.83	2375.5
	(47.22)	(87.21)
P3	386.6	2185.6
	(47.46)	(393.62)
P4	387.8	2387.8
	(22.30)	(99.09)

4.2.4 Kensiu /e/

Three words were used to elicit /e/ and these were *hubet* [hubet] 'medicine', *tenapes* [tənapes] 'sieve' and *ages* [ages] 'mosquitoes' (see Table 3.2). The frequencies of each word were equal for this vowel. Table 4.8 provides the frequencies of /e/ elicited from all five speakers.

Table 4.8. Frequencies of /e/ in each word

Kensiu word	Meaning	Frequency
[hubet]	medicine	10
[tənapes]	sieve	10
[ages]	mosquitoes	10

Figure 4.5 illustrates the distribution of /e/ in Kensiu. The distribution of /e/ was dispersed from front to central and close to mid-open in the vowel space. In short, the productions of these vowels were more widely distributed. The variation of the pronunciations elicited by each speaker of a language that is not used frequently and

possibly considered as an endangered language. As depicted from Figure 4.5, P3 produced /e/ with a great level of variation especially in relation to vowel fronting where the vowels were dispersed from front to the centre. There is also a considerable range for F2 values produced by P3 from 1658 Hz to 2652 Hz. The tokens of /e/ produced by P5 showed low variation in which all of them were clustered near each other. From Figure 4.5, it is still can be seen that P5 produced a more fronted /e/ than the rest. However, this time, the realisations of these vowels by P5 were lower than those produced by the other four speakers. Meanwhile, P4 appears to have produced this vowel higher than the rest although there were few instances of P3 located higher than P4. However, with only two instances of P3 being higher than P4, it cannot be concluded P3's /e/ were higher. Since most of the vowel /e/ produced by P3 were located further back and lower, further inspection for *tenapes* [tənapes] 'sieve' of the first and second elicitation were conducted as these two words were located at the close and front position. Based on the audio inspection, it was found that the pronunciation of the word *tenapes* by P3 almost resembles *tenapis* as in /i/. Figure 4.5 shows the distribution of /e/ in the vowel space.

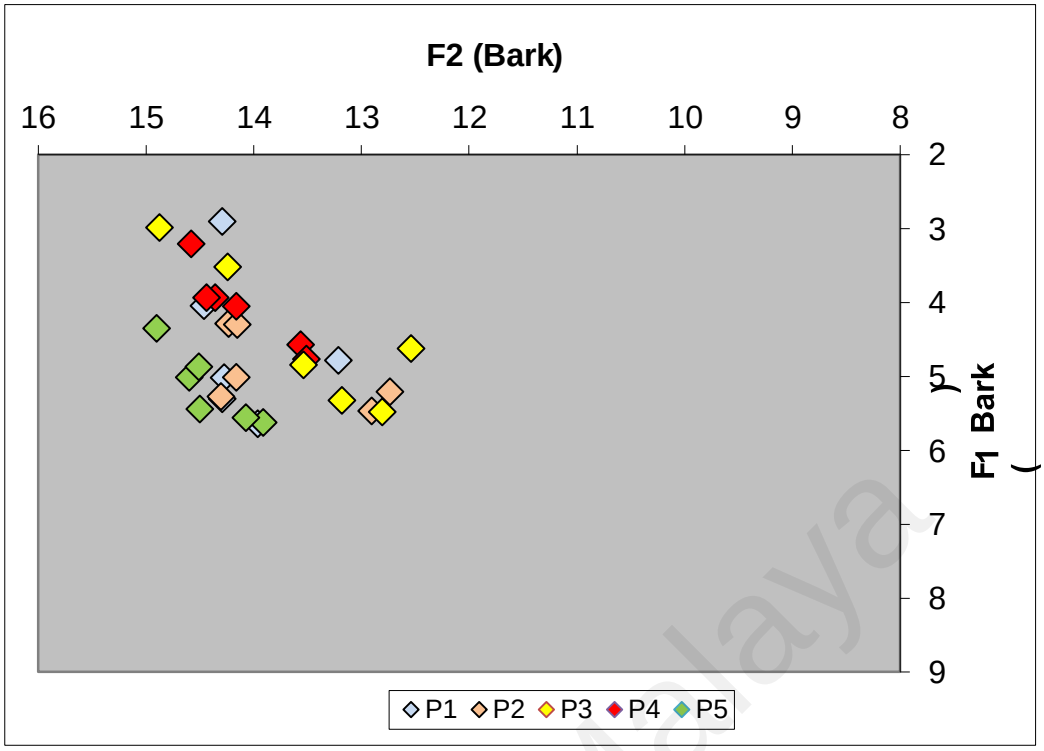


Figure 4.5. Scatter plot of /e/ in Kensiu

A one-way ANOVA showed that there were no significant differences between the average values of F1 from five speakers, $F(4, 25) = 1.86$, $p = 0.15$. This suggests that in terms of vowel height, all the speakers produced the /e/ in a similar way. A oneway ANOVA between the F2 means from the speakers also showed no significant

differences, $F(4, 25) = 2.08$, $p = 0.11$. This indicates that in terms of vowel fronting, the production of vowels between the speakers were also the same. Table 4.9 provides the formant measurements for /e/ produced by each speaker.

Table 4.9. Formant measurements for Kensiu /e/ in each speaker

Speaker	Ave. F1 and SD (Hz)	Ave. F2 and SD (Hz)
P1	487.83 (112.25)	2340.5 (162.28)
P2	523.5 (59.23)	2224 (243.32)
P3	472.17 (112.66)	2090.67 (391.49)

P4	427.00 (61.46)		2344.67 (170.00)
P5	549.00 (57.51)	2464.83 (144.00)	

4.2.5 Kensiu /ɛ/

The tokens for were extracted from *gehit* [gɛhit] 'bitter', *klapeh* [klapɛh] 'arm', *lebeh* [ləbɛh] 'bamboo', *bekes* [bəkɛs] 'food container', *bejakes* [bəjakɛs] 'adult', and *ges* [gɛs] 'gas'(see Table 3.2). Table 4.10 shows the frequency of /ɛ/ in each word elicited from the speakers.

Table 4.10. Frequencies of /ɛ/ in each word

Kensiu word	Meaning	Frequency
[gɛhit]	bitter	7
[klapɛh]	arm	8
[ləbɛh]	bamboo	9
[bəkɛs]	food container	6
[bəjakɛs]	adult	10
[gɛs]	gas	10

Figure 4.6 presents the scatter plot for /ɛ/. The vowels were scattered from close to mid open and front to the central position. The same case as /e/, there was little overlapping among the productions of these vowels by the five speakers. The productions of these vowels were mostly concentrated at the central and lower part of the vowel space. The unstable form of /ɛ/ in Kensiu suggests that all the speakers produced the vowel with a high degree of variation. As can be observed from Figure 4.6, speaker P4 kept producing higher vowels than the rest. Meanwhile, the vowels produced by P5 were scattered at the front area of the vowel space. It is noticeable that there were two clusters of the vowel

produced P5 clustered to each other. The first group was located higher and more fronted than the second group which was located slightly lower and back than the first one. Further inspection was carried out on the words produced by P5. There were four instances of P5 were positioned slightly higher which were *gehit* [gɛhit] 'bitter' and *bekes* [bəkɛs] 'food container' for the first and second elicitations. The sounds were listened to again, and it was found that the pronunciations of the vowel /ɛ/ were not as low as the rest of instances for /ɛ/. The words that were located near the P5 productions (represented in green) of *gehit* [gɛhit]

'bitter' and *bekes* [bəkɛs] 'food container' which were two instances of P1 (represented in blue), one instance of P3 (represented in yellow) and two instances of P4 (represented in red) also came from the same words which were *gehit* [gɛhit] 'bitter' and *bekes* [bəkɛs] 'food container'. Thus, it can be surmised that the vowel in these words was closer to the rising /e/ which was produced higher and more fronted. There was one outlier by P3 that was located further back the rest. That was the first elicitation for the word *lebeh* [ləbɛh] 'bamboo' (note that the second elicitation was removed due to the noise in the recording). Figure 4.6 illustrates the dispersion of /ɛ/ in the vowel space.

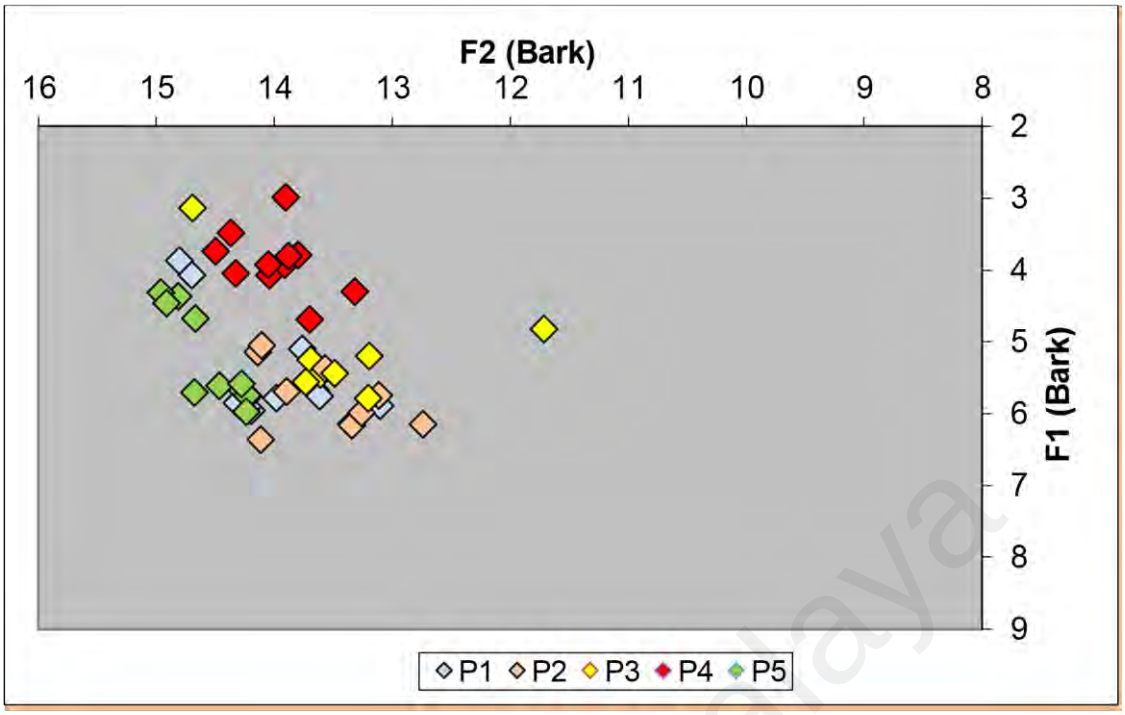


Figure 4.6. Scatter plot of /ɛ/ in Kensiu

A one-way ANOVA showed that there were significant differences between the F1 average values of five speakers, $F(4, 45) = 11.91$, $p < 0.05$. Tukey's HSD test of five speakers show that the F1 means were also significantly different between P1 ($M = 586.73$ Hz, $SD = 91.07$) and P4 ($M = 405.5$ Hz, $SD = 50.30$), P2 ($M = 620.33$ Hz, $SD =$ 57.19) and P4 ($M = 405.5$ Hz, $SD = 50.30$), P3 ($M = 544.13$ Hz, $SD = 94.24$) and P4 ($M = 405.5$ Hz, $SD = 50.30$), and P4 ($M = 405.5$ Hz, $SD = 50.30$) and P5 ($M = 567.75$ Hz, $SD = 75.10$). As F1 corresponds to the vowel height, this statistical analysis suggests that the speakers produced the vowel rather differently from each other.

In contrast, for F2 means of the five speakers, a one-way ANOVA showed that there were no significant differences, $F(4, 45) = 1.39$, $p = 0.25$. In terms of vowel fronting, this suggests that these vowels produced by the speakers in a similar manner.

Table 4.11 was presented to show the formant measurements of /ɛ/ in each speaker.

Table 4.11. Formant measurements for Kensiu /ɛ/ in each speaker

Speaker	Ave. F1 and SD (Hz)	Ave. F2 and SD (Hz)
P1	586.73	2303.55
	(91.07)	(198.02)
P2	620.33	2163.00
	(57.19)	(170.22)
P3	544.13	2101.00
	(94.24)	(300.27)
P4	405.5	2301.10
	(50.30)	(128.62)
P5	567.75	2374.42
	(75.10)	(477.26)

4.3 Vowel conflation

The following sections will present findings on the extent to which particular vowels are produced similarly.

4.3.1 Vowel contrast between /i/ and /ɪ/

Figure 4.7 shows the distribution of /i/ and /ɪ/ based on their average F1 and F2 values. There is considerable overlap between these vowels as shown in Figure 4.7. This indicates a lack of contrast among the vowels produced by the speakers suggesting that in Kensiu /i/ and /ɪ/ may be produced similarly. A two-tailed independent *t-test* was carried out to compare the F1 and F2 average values for both /i/ and /ɪ/. As anticipated, no significant differences were found between the average F1 for /i/ (M = 358 Hz, SD = 40) and /ɪ/ (M = 359 Hz, SD = 41); $t(67) = 0.06$, $p = 0.95$; $d = 0.02$. However, it should be noted that the effect size here is extremely small.

Likewise, there were no significant differences found in the average values of F2 for /i/ (M = 2638 Hz, SD = 238) and /ɪ/ (M = 2583 Hz, SD = 322); $t(53) = 0.88$, $p = 0.38$; $d = 0.35$. The effect size was between small to medium. Thus, there is some evidence that the vowels were conflated as the speakers produced the vowels similarly in terms of vowel height and vowel fronting. This could be due to the influence of Malay language where the community uses Kedah Malay quite predominantly in their life (see 3.2).

Figure 4.7 shows the comparisons of /i/ and /ɪ/ in the vowel space.

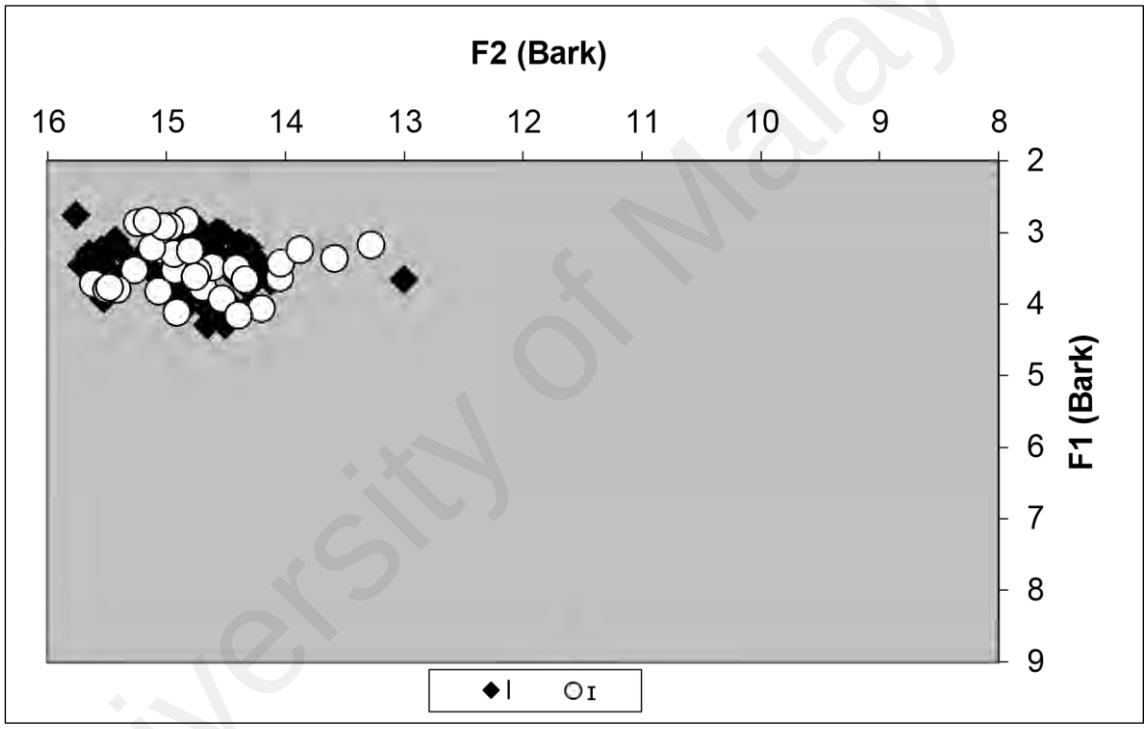


Figure 4.7. Scatter plot of /i/ and /ɪ/ in Kensiu

4.3.2 Vowel contrast between /ẹ/, /e/, /ɛ/

Figure 4.8 presents the distributions of /ɛ/, /ẹ/ and /e/ in the vowel chart. As discussed by Bishop (1996), /ẹ/ and /e/ are distinct vowels as they are located at different positions. The vowel /e/ is located at the mid-front, lower and more back than /ẹ/. A t-test was conducted to compare the values of F1 and F2 between the vowels. In terms of vowel height, a significant difference was found in the average of F1 values for /e/ (M = 492 Hz, SD = 90) and /ẹ/ (M = 422 Hz, SD = 44); $t(43) = 3.78$ $p < 0.05$; d =

0.99. The effect size exceeds Cohen's convention for a large effect which is 0.80.

No significant differences were found in terms of vowel fronting for average of F2 values for /e/ (M = 2293 Hz, SD = 257) and /ẹ/ (M = 2362 Hz, SD = 197); $t(53) = 1.14$, $p = 0.26$; $d = 0.30$ (between small to medium effect size, 0.20 and 0.50). These results indicate that there was no significant difference in vowel fronting in these vowels. Instead, the difference was found in terms of vowel height.

In terms of the placement for /e/ and /ɛ/ in the vowel chart, both vowels show low overlapping. The distributions between these vowels were almost the same. Based on the scatter plot in Figure 4.8, it can be seen that these vowels are scattered between close to mid-open and mid-front to mid-back. Based on a *t-test*, a significant difference was found in the average F1 values for /e/ (M = 492 Hz, SD = 90) and /ɛ/ (M = 545 Hz, SD = 104); $t(68) = -2.42$, $p < 0.05$, $d = 0.5$ (medium effect size). However, no significant difference was found in the average of F2 values for /e/ (M = 2293 Hz, SD = 257) and /ɛ/ (M = 2262 Hz, SD = 299); $t(68) = 0.4$, $p = 0.71$; $d = 0.11$ (extremely small effect size). This is consistent with Figure 4.8 where the difference between /e/ and /ɛ/

lies in vowel height instead of vowel fronting.

The distribution of /ɛ/ is contrasts with /ẹ/ as the former shows inconsistency among speakers while the distribution of /ẹ/ has a more consist distribution at the mid and mid-front position. However, commenting on the distribution of /ẹ/ and /ɛ/, Bishop (1996, p. 229) stated that "the /ɛ/ distribution and frequency are almost identical to that of /ẹ/". A *t-test* was carried out to compare the average values of F1 and F2 for both /ẹ/ and /ɛ/. A significant difference was found in the average values of F1 for /ɛ/ (M = 545 Hz, SD = 104) and /ẹ/ (M = 422 Hz, SD = 44); $t(72) = 7.24$, $p < 0.05$; $d = 1.54$. The effect size here exceeds Cohen's convention for a large effect, 0.80.However, there was no significant difference found in the average values of F2 for /ɛ/ (M = 2262 Hz, SD = 299) and /ẹ/ (M = 2362 Hz, SD = 197); $t(70) = -1.74$, $p = 0.11$; $d = 0.40$ (between small and medium effect size). This indicates that the distribution of /ẹ/ and /ɛ/ is the same in terms of the vowel fronting. However, they are very different in terms of vowel height.

Figure 4.8 presents the scatter plot of three vowels /ẹ/, /e/ and /ɛ/ in the vowel space.

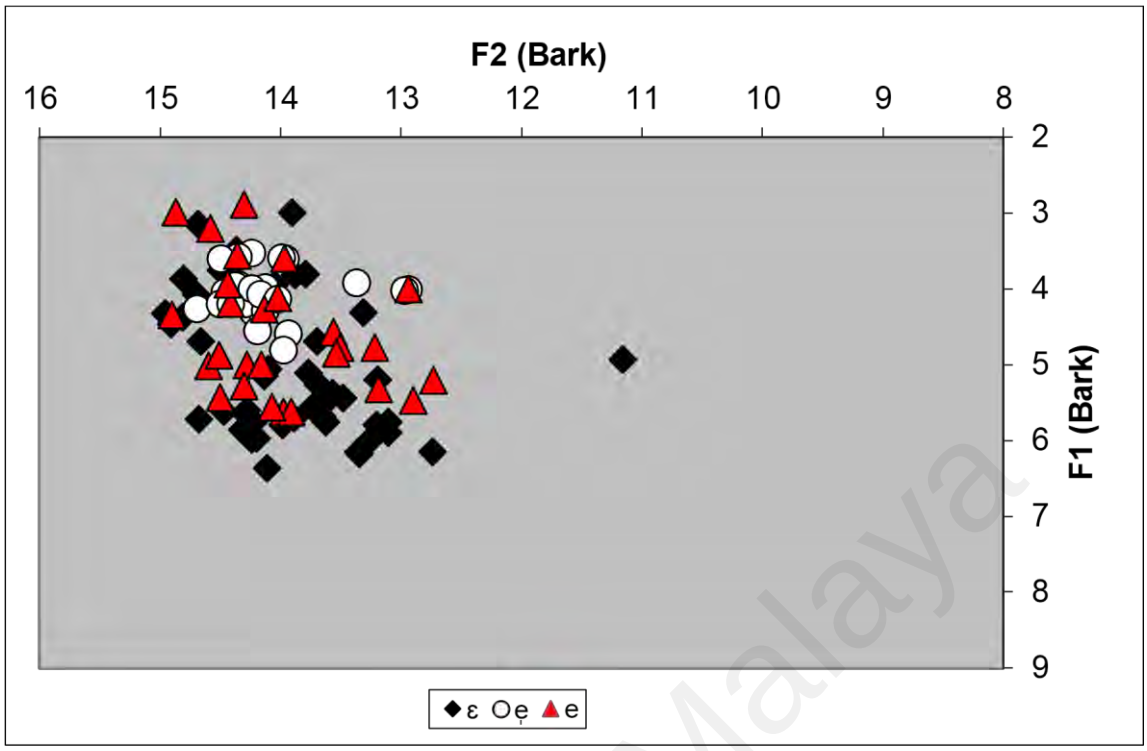


Figure 4.8. Scatter plot /ẹ/, /e/, /ɛ/ in Kensiu

4.4 Comparison of Kensiu and Kedah Malay

In the following sections, the findings of Kensiu are compared to Kedah Malay. The comparisons were based on Afiqah Jazmin's (2017) study. It is acknowledged that since this comparison is with another set of data, the results should be treated with caution.

4.4.1 Comparison of /i/ in Kensiu and Kedah Malay

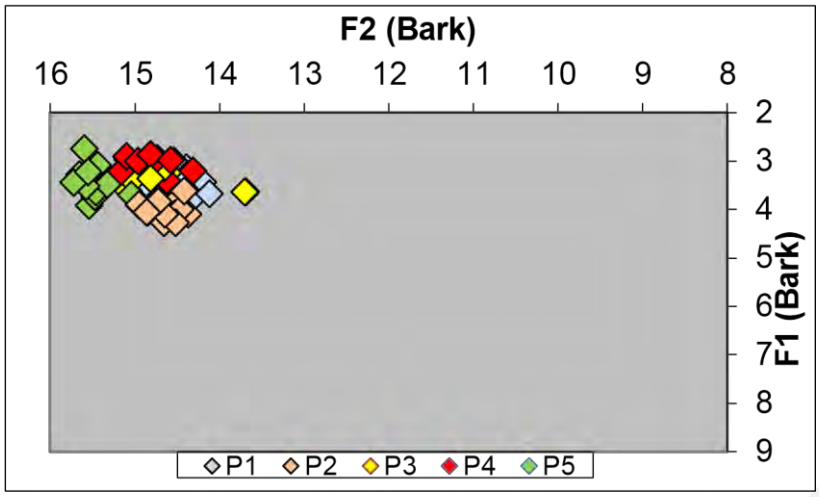
As shown in Figure 4.2, /i/ in Kensiu is located at the close and front position in the vowel space. The position of /i/ is similar to what was found by Afiqah Jazmin (2017) in Kedah Malay. The only difference was that, as shown in Figure 4.9, the distribution of /i/ shows low overlapping in the KM speakers compared to Kensiu which shows more overlaps among the production of the vowel. It is noticeable that KM speakers seemed to be slightly dispersed to the back meaning that there were some variations in the production

of /i/ in terms of vowel fronting. A comparison of the average values for F1 and F2 of Kensiu and KM indicates that Kensiu speakers produced /i/ almost similarly with KM speakers as their F1 and F2 values did not differ much. However, there was a relatively big difference in their F2 standard deviation between Kensiu and KM. That explains the huge dispersion in /i/ produced by KM. The table of measurements of /i/ for F1 and F2 in Kensiu and Kedah Malay and the scatter plot for /i/ for KM are presented in Table 4.12.

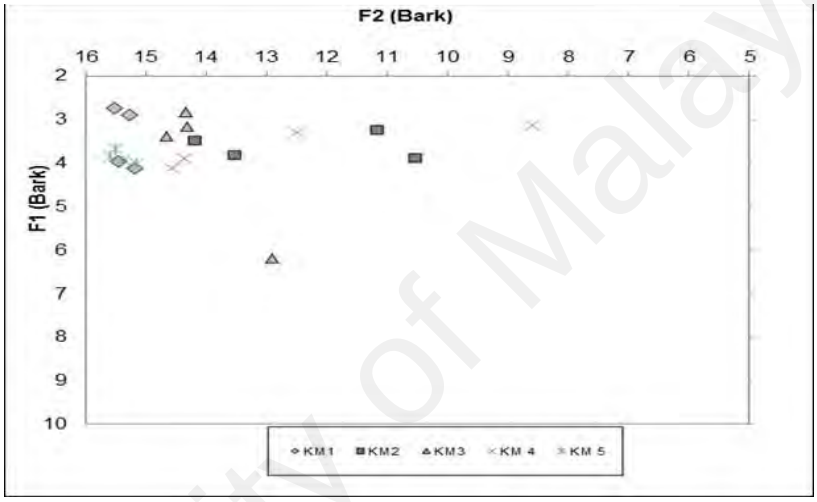
Table 4.12. Comparison of /i/ of F1 and F2 average values for Kensiu and KM

/i/	Ave F1 and SD (Hz)	Ave F2 and SD (Hz)	Ave F1 Bark	Ave F2 Bark
Kensiu	358.56 (39.71)	2637.95 (237.81)	3.46	14.84
KM**	382** (22.41)**	2381** (501.3)**	3.68**	14.07**

**taken from Afiqah Jazmin (2017, p. 42)



(a) Kensiu monophthongs



(b) Kedah Malay monophthongs**reproduced from Afiqah Jazmin (2017, p. 47)

Figure 4.9. Comparison of scatter plots for /i/ for Kensiu and KM monophthongs
**reproduced from Afiqah Jazmin (2017, p. 47)

4.4.2 Comparison of /e/ in Kensiu and KM

Figure 4.10 displays the vowel distribution of /e/ in KM. It can be seen that there is less variation of /e/ produced by the KM speakers meaning that they produced the same variation of /e/. From Figure 4.10, it can be seen that the /e/ from KM speakers were clustered at the front and mid-close position. Meanwhile, in Kensiu, these vowels were scattered from close to close to mid-open and front to central. Thus, it can be concluded that there was more stability in the production of /e/ in KM compared to Kensiu /e/ which was more dispersed. Based on the measurements for both Kensiu and KM, it clearly shows

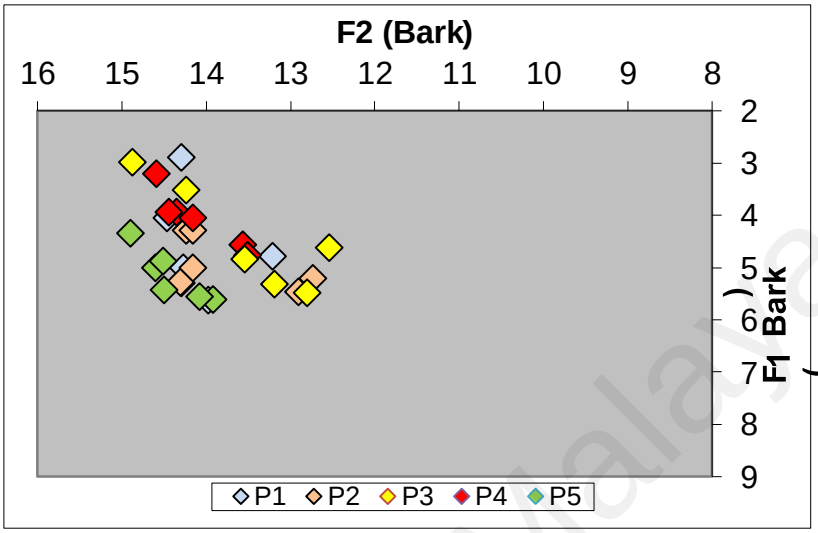
that in terms of vowel height, Kensiu speakers produced /e/ more spread out in the vowel space as it has a higher standard deviation (89.58) than KM speakers (24.89). The table of measurements of /e/ for F1 and F2 in Kensiu and Kedah

Malay and the scatter plot for /e/ for KM are presented in Table 4.13.

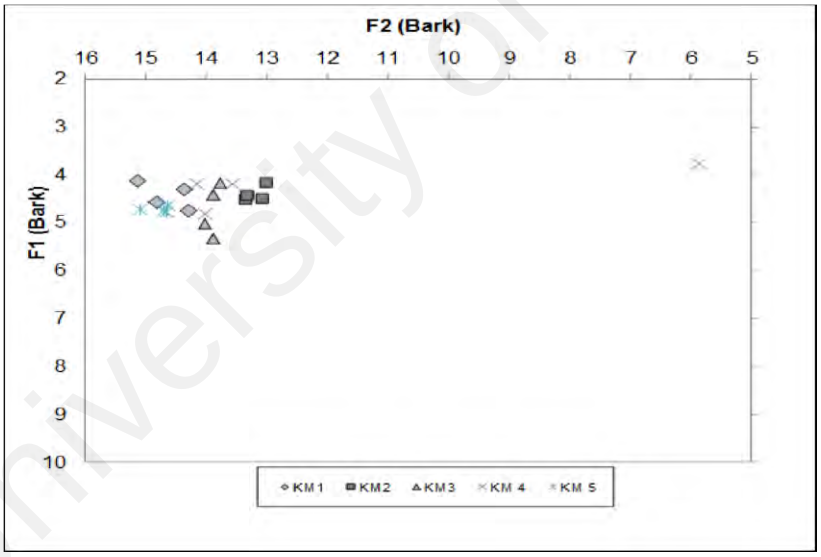
Table 4.13. Comparison of /e/ of F1 and F2 average values for Kensiu and KM

/e/	Ave F1 and SD (Hz)	Ave F2 and SD (Hz)	Ave F1 Bark	Ave F2 Bark
Kensiu	491.90 (89.58)	2293.10 (257.47)	4.67	13.97
KM**	474** (24.89)**	2262** (324.17)**	4.51**	13.82**

**taken from Afiqah Jazmin (2017, p. 42)



(a)Kensiu monophthongs



(b) Kedah Malay monophthongs**reproduced from Afiqah Jazmin (2017, p. 48)

Figure 4.10. Comparison of scatter plots for /e/ for Kensiu and KM monophthongs

4.4.3 Comparison of /ɛ/ in Kensiu and KM

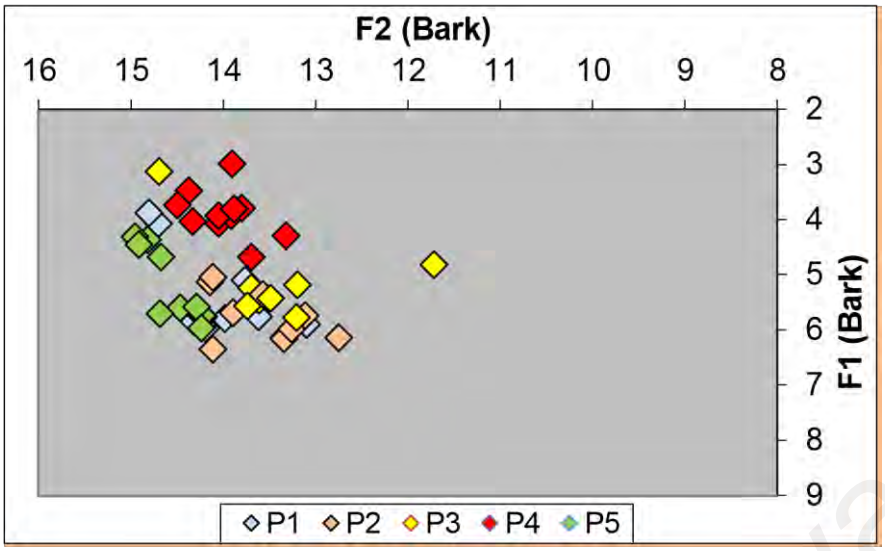
Figure 4.11 illustrates the distribution of /ɛ/ in KM. The distribution of /ɛ/ in KM was scattered at mid-close and front position and slightly moving to the centre. For the vowel height, these were slightly different compared to the vowels in Kensiu as Kensiu

speakers produced a higher /ɛ/ at the close position and moving lower to the mid-open. Once again, Kensiu shows more instability in the production of this vowel compared to KM which shows consistency between the speakers. A comparison of the values of the standard deviations shows that Kensiu was more spread in terms of vowel height, but KM had a higher standard deviation in F2. The table of measurements of /e/ for F1 and F2 in Kensiu and Kedah Malay and the scatter plot for /ɛ/ for KM is presented in Table 4.14.

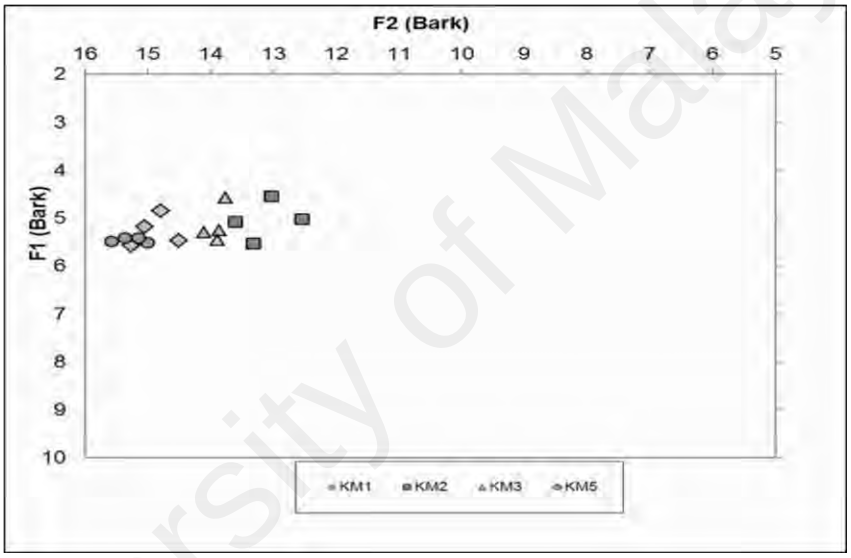
Table 4.14. Comparison of /ɛ/ of F1 and F2 average values for Kensiu and KM

/ɛ/	Ave F1 and SD (Hz)	Ave F2 and SD (Hz)	Ave F1 Bark	Ave F2 Bark
Kensiu	545.16 (103.72)	2262.36 (299.35)	5.12	13.89
KM**	529** (66.41)**	2234** (579.33)**	4.98**	13.57**

**taken from Afiqah Jazmin (2017, p. 42)



(a) Kensiu monophthongs



(b) Kedah Malay monophthongs**reproduced from Afiqah Jazmin (2017, p. 50)

Figure 4.11. Comparison of scatter plots for /ɛ/ for Kensiu and Kedah Malay monophthongs

4.5 Comparison of Kensiu and Standard Malay

Since Kensiu speakers also appear to speak a less dialectal form of Malay (which is closer to the spoken form in the Central and Southern west coast of Peninsular Malaysia) with non-Kedah speakers, a comparison with the spoken SM was done. The comparison was based on Yusuf's (2013) study which there was six SM vowels /i/, /e/, /a/, /u/, /o/ and /ə/. As predicted, there were similarities between the Kensiu and Standard Malay specifically in the position of the vowels. The vowel /i/ for SM was located at the

close front position consistent with Kensiu. The positions of /e/ in Kensiu were lower and slightly back than /i/, and this is similar with Yusuf (2013). The difference between Kensiu and SM was the variation of /ə/ and /ɛ/. In this study on

Kensiu, the vowel /ə/ was not analysed as it is considered as the central vowel.

However, the vowel /ɛ/ does not exist in SM. Table 4.15 shows the average values of F1

and F2 for both Kensiu and SM.

Table 4.15. Comparison of /i/ and /e/ average values for Kensiu and SM

	Kensiu		Standard Malay	
	Ave F1 and SD (Hz)	**Ave F2 and SD (Hz)**	**Ave F1 and SD (Hz)**	**Ave F2 and SD (Hz)**
/i/	358.56	2637.95	428**	2703**
	(39.71)	(237.81)	(18.58)**	(90.55)**
/e/	491.90	474	567**	2417**
	(89.58)	(24.89)	(18.97)**	(122.19)**

**taken from Yusuf (2013, p. 274)

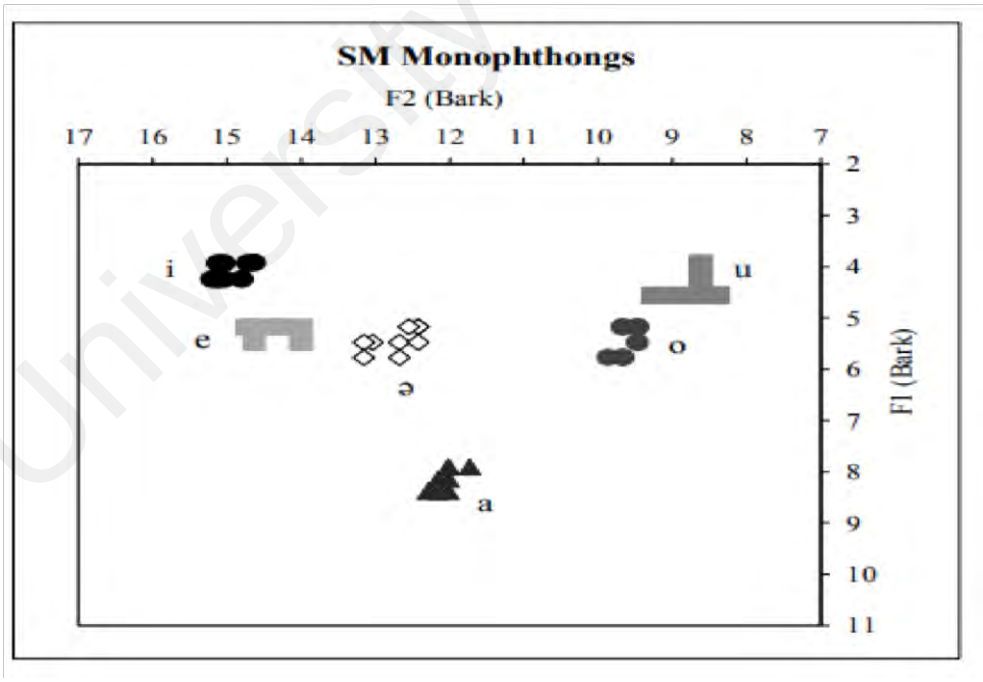


Figure 4.12. Scatter plot for SM monophthongs
**reproduced from Yusuf (2013, p. 274)

4.6 Summary

The findings showed the presence of four oral front vowels in Kensiu which were /i/, /ẹ/, /e/, /ɛ/. Contrary to what has been discussed previously by Bishop (1996), there was no distinction between /i/ and /ɪ/ in Kensiu Kedah. The production of /i/ and

/ɪ/ were at close and front positions. Similar findings were reported in Pillai, Zuraidah Mohd. Don, Knowles and Tang (2010) that /ɪ/ and /i:/ are not distinguished among Malay speakers. The vowel inventory of Kensiu was almost similar to the vowel inventory of Kedah Malay where there was an existence of /ɛ/. However, further analysis should be carried out with /ẹ/ to make sure that variety still exists or have merged into /e/. There may be a possibility that /ẹ/ had merged with /e/ but from the analysis, it was found that /ẹ/ and /e/ only differed in terms of vowel height although predictably, these vowels were produced similarly in terms of vowel fronting (see 4.3.2). This needs to be examined further as the tokens collected were not sufficient to ascertain this pattern. Furthermore, further investigations need to be done for the comparisons between Kensiu with Kedah Malay and SM as the data from those studies were obtained from different speakers which have different speaking environments. It must also be noted that the studies by Afiqah Jazmin (2017) and Yusuf (2013) used a different methods to collect their data. However, it can be assumed that Kensiu speakers produced the same vowel quality as Kedah Malay and Standard Malay as there were only slight differences in terms of the vowel quality.

In terms of the location of /i/, it is similar with Kedah Malay and SM as the majority of the distributions were scattered at the close and front position. However, the distribution of /e/ and /ɛ/ in Kensiu are slightly different in comparison to Kedah Malay as both vowels were scattered from the higher which is at the close position. The speakers showed great variation and instability when producing these vowels. As discussed in (see 4.2.4), this is normal for the endangered languages and is similar to what has been reported

for Malacca Portuguese (Pillai, Siti Raihan, Wan Aslynn, Roshidah Hassan & Phillip, 2019). In general, speaker P5 produced more fronted vowels in all vowels and P4 had a tendency to produce higher vowels than the other speakers.

CHAPTER 5

CONCLUSION

5.1 Introduction

This chapter summarises the important findings in relation to the research questions of this study. This study sought to examine the characteristics of the front oral monophthongs in Kensiu based on the first (F1) and second formant (F2) measurements of the vowels and to find out if there are vowel conflations between these front oral monophthongs. Besides that, the limitations of this study will be explained in this chapter. At the end of this chapter, recommendations for future studies are provided.

5.2 Summary of the study

The present study was designed to examine the acoustic properties of front monophthong vowels in Kensiu. The findings are divided into two sections as followed.

5.2.1 Research question 1

What are the characteristics of the front oral monophthongs in Kensiu based on the first (F1) and second formant frequencies (F2) measurements of the vowels?

It was found that only four front vowels in Kensiu contrary to what Bishop (1996) has discussed which were /i/, /ẹ/, /e/, /ɛ/. The formant values of /ɪ/ were similar to /i/ and thus we can assume that in terms of vowel quality, there is one rather than two high front vowels in Kensiu. For the placement of vowel /i/, the majority of the speakers produced higher and fronted /i/. The same thing applied to /ɪ/ where it is positioned at the close and fronted location in the vowel space. In the production of /ẹ/, there were overlaps at the mid-close and mid-front position. Next, the position of vowel /e/ appears to be unstable where the productions of the vowels were scattered from front to central and close to mid-open positions in the vowel space. It was difficult to ascertain the exact position of /e/ as they were widely distributed in the vowel space. In a similar vein, the

realisations of /ɛ/ were widely distributed from front to central and close to mid-open. In general, the Kensiu vowel inventory is almost the same as Kedah Malay especially as they both contain the vowel /ɛ/.

5.2.2 Research question 2

To what extent is there vowel conflation between these front oral monophthongs?

As mentioned in the previous section the vowel inventory of Kensiu in Baling has four rather than the five vowels described for the variety spoken in Thailand (Bishop, 1996). The vowel /ɪ/ has merged into /i/ with no significant differences in terms of vowel height and vowel fronting. For the vowel contrast between /ẹ/, /e/ and /ɛ/, these vowels show no significant differences in terms of their vowel fronting as expected, but there were significant differences in their vowel height. This suggests that the difference between these vowels lies in their vowel height, and they appear to be contrasted. However, this should be further analysed as the number of tokens collected from each vowel were different as this may somehow affect the results.

5.3 Limitations of the study

In this study, only one group of participants consists of five native speakers were employed. Thus, in terms of the sample size, it was limited to five speakers only. This study was also limited by the absence of the male participants as all the five participants were female. The age of the participants is 40-60 years old. Hence, the findings should not be generalised to all age groups. The scope of this study was also restricted in terms of the production of vowels as it was narrowed down to examine only front oral vowels. In spite of this limitation, this study adds to our understanding of the acoustic properties of oral vowels in Kensiu. The number of target words was also limited as it was solely based on the glossary from Bishop (1996). It did not include the words which were not in

the glossary although the words were spoken from the participants and it followed the plosives and fricatives environments.

5.4 Further research

Further studies should be done to examine the other vowels in Kensiu to build a more complete picture of its vowel inventory. In addition, the consonants and prosodic features should also be looked into if there is to be a more complete inventory and description of the Kensiu sound system. As discussed in Chapter 4 (see 4.6), further analysis on /ẹ/ should be done as there is a possibility that /ẹ/ does not exist in Kensiu. However, in comparison to /e/ and /ɛ/, /ẹ/ shows high stability. The findings on /e/ and

/ɛ show some degree of instability that there were variation in the vowel quality among the speakers. Thus, thorough inspection on the speakers is needed to determine whether those vowels are influenced by other languages such as Malay or if it is normal for endangered language like Kensiu to have the instability of vowels.

Future research should be done with different age groups such as younger participants and it is also suggested that the male gender is included for the study as well. Such focus on different age groups will reveal whether younger participants are still fluent in their native language. Last but not least, future research should be more cautious in deciding the target words as it may affect the number of tokens elicited by the speakers. This also will affect the data, as well as the number of tokens for each vowel, is not balanced. Detailed preparation on the materials should be done thoroughly before the data collection begins.

5.5 Conclusion

In conclusion, the findings from this study have significant implications for understanding the acoustic properties of Kensiu vowels. It also helps to add the knowledge of Kensiu in Kedah as previous research was done in Kensiu which is located in Yala, Thailand. Other than that, this study suggests that further research should focus on determining consonants and prosodic features as well as it would be great help to have more complete inventory and the description of Kensiu sound system. The recommendations and suggestions for future research would provide a great help in preserving the Kensiu language in the future. It also helps in understanding Kensiu language as a whole.

REFERENCES

Afiqah Jazmin, A. (2017). *An acoustics study of the monophthongs of Kedah Malay* (Unpublished master's dissertation). University of Malaya, Kuala Lumpur.

Alias Abd Ghani & Salasiah Che Lah. (2015). The Semang Kensiu Orang Asli of Lubuk Legong, Baling: Their language and cultural endangerment. *Procedia Social and Behavioral Sciences, 208*, 21-30. doi:10.1016/j.sbspro.2015.11.177

Asmah Haji Omar. (1991). *Kepelbagaian fonologi dialek-dialek Melayu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Baxter, A. N., & De Silva, P. (2004). *A dictionary of Kristang:(Malacca Creole Portuguese*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Beck, D., & Lam, Y. (2008). Language loss and linguistic suicide: A case study from the Sierra Norte de Puebla, Mexico. *Toronto Working Papers in Linguistics 27*, 516.

Benjamin, G. (1976). Austroasiatic subgroupings and prehistory in the Malay Peninsula. *Oceanic Linguistics Special Publications*, *13*, 37-128.

Benjamin, G. (2012a). The Aslian languages of Malaysia and Thailand: An assessment. *Language Documentation and Description*, *11*, 136-230.

Benjamin, G. (2012b). Why you should study Aslian languages. Paper presented at the Second International Conference on Linguistics, Literature and Culture, Penang, Malaysia 7-9 November 2012. 1-35.

Bishop, N. (1996). A preliminary description of Kensiw (Maniq) phonology. *Monkhmer Studies*, 25, 227-253.

Blust, R. (2013). *The Austronesian languages*. Canbera: Australian National University.

Boersma, P & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.08, retrieved 1 September 2019 from http://www.praat.org/

Burenhult, N. (2001). Linguistic aspects of the Semang. In Rashid, R., & Karim, W. J. (Eds.), *Minority cultures of Peninsular Malaysia survivals of indigenous heritage* (pp. 75-80). Penang: The Academy of Social Sciences (AKASS).

Burenhult, N., Kruspe, N., & Dunn, M. (2011). Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula. In N. J. Enfield (Ed.), *Dynamics of human diversity: The case of mainland Southeast Asia* (257275). Canberra: Pacific Linguistics.

Butcher, A. (2013). Research methods in phonetic fieldwork. In *The Bloomsbury companion to phonetics* (pp. 57-78). London/ New York: Bloomsbury Academic.

Carey, I. (1970). The Kensiu Negritos of Baling, Kedah. *Journal of the Malaysian Branch of the Royal Asiatic Society*, *43*(217), 143-154.

Coluzzi, P. (2017). The vitality of minority languages in Malaysia. *Oceanic Linguistics, 56*(1), 210-225.

Coluzzi, P., Riget, P. N., & Xiaomei, W. (2017). Language vitality among the Orang Asli of Malaysia: The case of the Mah Meri on Telo'Gunjeng (Carey Island, Selangor). *International Journal of the Sociology of Language*, *2017*(244), 137-161.

Coupe, A. R. (2014). Strategies for analyzing tone languages. *Language Documentation & Conservation*, *8*, 462-489.

Crawford, J. (1995). Endangered Native American languages: What is to be done, and why? *The Bilingual Research Journal* Winter, *19*(1), 17-38.

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative and qualitative research* (3rd ed). Upper Saddle River, NJ : Pearson/Merrill Prentice Hall.

Department of Orang Asli development. (2018). Retrieved January 27, 2020, from https://www.jakoa.gov.my/

Department of Statistics Malaysia. (2015). Retrieved from https://www.dosm.gov.my/v1/

Derdemezis, E., Vorperian, H. K., Kent, R. D., Fourakis, M., Reinicke, E. L., & Bolt, D. M. (2016). Optimizing vowel formant measurements in four acoustic analysis systems for diverse speaker groups. *American Journal of SpeechLanguage Pathology*, *25(3),* 335-354.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, *27*(1-2), 47-55.

Deterding, D. (2003). An instrumental study of the monophthong vowels of Singapore English. *English World-Wide*, *24*(1), 1-16.

Dowd, A., Smith, J., & Wolfe, J. (1998). Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time. *Language and Speech*, *41*(1), 1-20.

Duangchan, P. (2006). Glossary of Aslian languages: The northern Aslian languages of Southern Thailand. *Mon-Khmer Studies*, *36*, 207-224.

Dunn, M., Burenhult, N., Kruspe, N., Tufvesson, S., & Becker, N. (2011). Aslian linguistic prehistory: A case study in computational phylogenetics. *Diachronica*, *28*(3), 291-323.

Gold, E., & Earnshaw, K. (2019). San Diego accents: A preliminary investigation into the vowel spaces of young males and females. *The Journal of the Acoustical Society of America*, *146*(4), 3009-3009.

Gordon, M. (2003). Collecting phonetic data on endangered languages. In M. J. Solé, D. Recasens & J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 207-210).

Grimes, B. F. (2000). *Ethnologue*. Dallas, TX : SIL International.

Heinrich, P., Miyara, S., & Shimoji, M. (2015). *Handbook of the Ryukyuan languages: history, structure, and use*. Berlin: Walter de Gruyter.

Hoffer, B. L. (2005). Language borrowing and the indices of adaptability and receptivity. *Intercultural communication studies*, *14*(2), 53.

Indirawati Zahid & Abdul Hamid Mahmood. (2016). *Pengantar fonetik*. Dewan Bahasa dan Pustaka, Kuala Lumpur.

Kent, R. D., & Read, C. (2002). *Acoustic analysis of speech, 2nd* edition. Thomson Learning.

Kruspe, N. (2004). *A grammar of Semelai*. Cambridge: Cambridge University Press.

Kruspe, N., & Hajek, J. (2009). Mah Meri. *Journal of the International Phonetic Association*, *39*(2), 241-248.

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. City?: Wiley-Blackwell.

Laird, P. (1979). Ritual, territory and region: the Temoq of Pahang, West Malaysia. *Social Analysis: The International Journal of Social and Cultural Practice*, *1*, 54-80.

Laman Web Rasmi Jabatan Kemajuan Orang Asli Malaysia. (n.d.). Retrieved December 19, 2019, from http://www.jakoa.gov.my/

Masron, T., Masami, F., & Norhasimah Ismail. (2013). Orang Asli in Peninsular Malaysia: population, spatial distribution and socio-economic condition. *Journal of Ritsumeikan Social Science Humanities*, *6*, 75-115.

Nagata, S. (2005). Subgroup 'names' of the Sakai (Thailand) and the Semang (Malaysia): a literature survey. *Anthropological Science*, *114*(1), 45-57.

Observation on the state of the indigenous human rights in Malaysia. (n.d.). Retrieved January 30, 2020, from https://www.culturalsurvival.org/sites/default/files/Malaysia 2018 UPR Report.pdf

Phillips, T. C. (2007). Brief phonology of Semai: Betau dialect. In Chong Shin, K. Harun & Y. Alas (Eds.), *Reflections in Southeast Asian seas. Essays in honour of Professor James T. Collins (Book II)*. Pontianak: STAIN Pontianak Press.

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

Rawlins, K. (2006). Stress and epenthesis in Mohawk. *Ms., UC Santa Cruz*.

Reid, L. A. (1994). Morphological evidence for Austric. *Oceanic linguistics*, 323-344.

Renganathan, S., & Kral, I. (2018). Digital preservation of language, cultural knowledge and traditions of the indigenous Semai. In S. L. Chong (Ed), *SHS Web of Conferences Vol 53 - International Conference on Humanities and Social Sciences (ICHSS 2018)* (pp. 1-5).

Rohani Mohd Yusof, & Nur Hidayah Mohamed Suleiman. (2014). Sistem Bilangan dalam Bahasa Orang Asli Duano. *Jurnal Bahasa, 14*(1), 45-59.

Roshidah Hassan, Kamila Ghazali & Asmah Haji Omar (2015). Vitality of the Orang Asli languages in Gerik, Perak. *Oceanic Linguistics*, *54*(2), 492-506.

Showkat, N., & Parveen, H. (2017). Non-probablity and probability sampling. Retrieved from file:///C:/Users/Win8/Downloads/150114863016cretextnayeemshowkatnonprobabilityandprobabilitysampling%20(2).pdf

Sidwell, P. (2009). *Classifying the Austroasiatic languages: history and state of the art*, vol. 76. Munich: Lincom Europa.

Simons, G. F., & Fennig, C. D. (Eds.). (2017). *Ethnologue: Languages of Asia*. Dallas, Texas: Sil International.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, *3*(2), 621-640.

Statistik Penduduk Masyarakat Orang Asli Mengikut Sub Etnik - MAMPU. (2018, December 12). Retrieved from http://www.data.gov.my/data/ms_MY/dataset/statistik-penduduk-masyarakatorang-asli-mengikut-sub-etnik

Stevens, M., Kruspe, N., & Hajek, J. (2006, May). Register in Mah Meri: A preliminary phonetic analysis. In R. Hoffmann & H. Mixdorff (Eds.), *Proceedings of Speech Prosody, 3rd International Conference, Dresden, May 2-5, 2006* (pp. 269-272).

Styler, W. (2017). Using Praat for linguistic research. *University of Colorado at Boulder Phonetics Lab*.

Tengku Intan Suzila & Teo, K. S. (2015, February). Batek Hep in the Malay world: Preliminary analysis of loanwords. In *Interdisciplinary Behavior and Social Sciences: Proceedings of the 3rd International Congress on Interdisciplinary Behavior and Social Science 2014 (ICIBSoS 2014), 1-2 November 2014, Bali, Indonesia.* (p. 203-207). CRC Press.

Teoh, B. S. (1988). *Aspects of Malay phonology revisited: a non-linear approach.* (Unpublished doctoral dissertation). University of Illinois, USA.

Teoh, B. S. (1994). *The sound system of Malay revisited*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Tunley, A. (1999). *Coarticulatory influences of liquids on vowels in English.* (Unpublished doctoral dissertation). University of Cambridge, England.

Wessendorf, K. (2008). *The Indigenous World 2008*. Copenhagen, Denmark: International Work Group for Indigenous Affairs.

Yager, J., & Burenhult, N. (2017). Jedek: A newly discovered Aslian variety of Malaysia. *Linguistic Typology, 21*(3), 493-545.

Yusuf, Y. Q. (2013). *A comparative study of vowels in the Acehnese language spoken in Kedah, Malaysia and Aceh, Indonesia*. (Unpublished doctoral dissertation). University of Malaya, Kuala Lumpur.

Yusuf, Y. Q., & Pillai, S. (2016). An instrumental study of oral vowels in the Kedah variety of Acehnese. *Language Sciences*, *54*, 14-25.

Zwicker, E., & Terhardt, E. (1980). Analytical expression for critical-band rate and critical bandwith as a function of frequency. *Journal of the Acoustical Society of America, 68(5),* 1523-1525.