PERCEPTION OF ENGLISH VOWEL CONTRASTS BY ACEHNESE-INDONESIAN BILINGUAL LEARNERS OF ENGLISH

First Received: (the date you send the manuscript) Final Proof Received: (the date you send the final version)

Abstract

Introduction: Previous studies have reported that L2 learners filter non-native L2 sounds through their existing native L1 sounds when learning L2 sounds. The degree of similarity between L1 and L2 sounds predicts the ease of acquisition of non-native L2 sounds. Indonesian learners of English from Aceh are likely to speak two languages before they learn English at school. Thus, they have larger phonemic inventories to rely on when learning English sounds. **Objective:** In this study we seek to learn how Acehnese-Indonesian bilinguals perceive five English vowel contrasts (/ɪ/ - /iː/, /æ/ - /ɛ/, /ʌ/ - /ɑː/, /ʊ/ - /uː/, and /ɑː/ - /ɔː/). Special interest is given to their perception of English vowel contrasts that are new, similar, and identical to Acehnese and Indonesian vowels. **Method:** A group of 31 high school students (*N=31*) from an Islamic boarding school in Aceh participated in this study. An AX test comprising repetition and minimal pairs of the English vowel contrasts in CVC word context were randomly presented to the students. **Findings:** The findings indicate that the Acehnese-Indonesian bilinguals discriminate the /æ/ - /ɜː/ and /ɑː/ - /ɔː/ vowel contrasts better than the /ɪ/ - /iː/, /ʌ/ - /ɑː/ and /ɑː/ - /ɔː/ vowel contrasts. The vowel contrast in which both vowels are novel to Acehnese and Indonesian was moderately discriminated compared to the pairs in which one vowel is similar to both Acehnese and Indonesian vowels. **Conclusion:** Students perceived pairs with one vowel similar and identical to Acehnese-Indonesian better than pairs with both vowels or one vowel new to Acehnese and Bahasa Indonesia.

Keywords: Acehnese-Indonesian bilinguals, perception of English vowel contrasts, phonemic contrast

INTRODUCTION

First language (L1) experience might impede the acquisition of non-native phonemes if either one or both vowels in the L2 vowel contrasts are realized differently or absent in the learner's L1 (Flege, 1995; Best, 1994; Iverson et al., 2003). English vowels are known to differ in terms of quality and length, while Indonesian and Acehnese vowels differ only in terms of quality. Some English vowels are absent from Indonesian and Acehnese vowel system such as /æ/, /ɪ/, /ʊ/, /ɑː/ and /ɜ˞/, some are identical such as /ʌ/ and /ɛ/, while others are similar but with shorter duration such as /iː/, /uː/, and /ɔː/. Acehnese-Indonesian bilinguals may face difficulties discriminating against these vowels if they fail to realize them differently. Previous studies indicate that Acehnese speakers tend to produce English vowel contrasts similarly when one of the vowels is absent in their vowel system. For example, Fata, Fitriani, Mohammad, and Yusuf (2017), found that Acehnese learners of English produced English vowels similarly for the following vowel pairs: /ɪ/ and /iː/, /e/ and /æ/, /u/ and /ʊ/, and /ɑː/ and /ɔ/. In terms of perception, Perwitasari (2018) reported that Javanese and Sundanese learners of English had difficulties perceiving English vowels that were new to their L1 system. They were found to have higher error rates perceiving the English vowels /ɪ/, /ʊ/, /ɔː/, /ɑː/, /ɜː/, and /æ/. It was also reported that Indonesian learners of English with Indonesian, Javanese, Sundanese, and Sulawesi as their respective first languages had difficulties discriminating English vowel contrasts in which one vowel new to their vowel system such as /ɪ/ - /iː/ /ɑː/ - /ɔː/ and /ʌ/ - /ɑ/ (Perwitasari, 2013).

Unlike Indonesian, Javanese and Sundanese, the Acehnese language has a bigger vowel inventory and has sounds that do not exist in Indonesian, Javanese, and Sundanese. Acehnese also has vowels that are similar to English, such as /ʌ/, /ɔ/, and /ɛ/. While Acehnese

/ʌ/ and /ɛ/ are identical to English /ʌ/ and /ɛ/, the vowel /ɔ/ in Acehnese is likely to be produced with a relatively shorter duration than English /ɔ:/. Iverson and Evan (2007) argue that speakers with larger phonemic inventories may be able to make use of their existing vowels and perceive non-native vowels better than speakers with smaller phonemic inventories. This is based on the findings of their research on the perception of the English vowels by the German, Norwegian, French and Spanish speakers. They found that the German and Norwegian speakers who have larger vowel systems were able to perceive English vowels better than the Spanish and French speakers who have smaller vowel systems. However, Elvin, Escudero, and Vasiliev (2014) found that size was not a good predictor for a more accurate perception of vowels. For instance, when discriminating non-native sounds of Brazilian Portuguese, Australian speakers were able to outperform the Iberian Spanish speakers, although the latter has larger vowel inventories. In fact, Alispahic, Mulk, and Escudero (2017) suggest that L1–L2 acoustic similarity rather than the size of the vowel inventory better predict the ease of non-native vowel perception.

The Speech Learning Model (SLM) posits that if two similar phonemes between the first (L1) and second language (L2) are not distinguished, a perceptual association between the sounds may be created and obstruct the construction of new phonemic categories in the L2 (Flege, 1995). In relation to this, the Perceptual Assimilation Model (PAM) further specifies that the ease of perception between a native and target sound can be predicted by the degree of similarity between the two sounds (Best, 1995). In addition, Escudero (2005), in her Second Language Linguistics Perception (L2LP) model, suggests that a direct link between perception and production exists, and comparisons of acoustic properties between L1 and L2 define listeners' ability to map and discriminate non-native sounds. These models are plausible as it has been found that adult second language learners tend to use their L1 to filter L2 speech (Bongaerts, van Summeren, Planken, & Schils, 1997; Escudero, 2005). The fact that they already have a particular phonemic inventory in their native language may encourage them to assimilate certain L2 sounds into the existing L1 sounds (Fabra, 2005).

Even though studies on speech perception of English learners learning the language in foreign language settings have been on the rise in the last decade (e.g., Alispahic, Mulak, & Escudero, 2017; Evans & Alshangiti, 2018), studies on the bilingual speakers of Southeast Asian languages especially Indonesian are scarce (Perwitasari, 2013; Fata et al., 2017; Perwitasari, 2018). Studies on the perception of English vowels by Indonesian listeners were conducted by Perwitasari (2013, 2018). Most listeners in these studies are Javanese, Sundanese and monolingual Indonesian speakers. These languages have smaller vowel systems compared to Acehnese and have less similar vowels to English. In relation to Acehnese speakers, studies like Fata et al. (2017) only focused on production. Thus, a study on the perception of English vowel contrasts by Indonesian bilinguals whose local language has more vowels similar to English would enable us to understand if their existing vowel systems help them discriminate English vowels. This study will contribute to our understanding of the way multilingual speakers with larger vowel inventories acquire English sounds. Considering that Indonesian speakers may have different vowel inventories depending on the languages that they speak, the results of the study could be used to develop English pronunciation course materials tailored specifically to Acehnese-Indonesian learners of English. The current study aims to fill the gap in the literature description of English vowel perception by Indonesian learners of English who are Acehnese-Bahasa Indonesia bilinguals. This study is also expected to contribute to the debate on emerging research in speech perception theories in terms of whether speakers with larger phonemic inventories and have more vowels similar to the target language are able to make use of their

existing vowels when discriminating non-native vowel contrasts. In relation to this, the following research questions are addressed in the current study:

1. To what extent do Acehnese-Indonesian bilinguals perceive vowel contrast in English
2. To what extent do their existing vowel systems contribute to their perception of English vowel contrast?

Language Perception Model

There have been at least three theoretical models proposed to describe the perceptual pattern of speech sound, Speech Learning Model (SLM), Perceptual Assimilation Model (PAM), and Second Language Linguistic Perception (L2LP). The acquisition of non-native vowels and consonants in natural linguistic contexts is explained in SLM (Flege, 1995, 1999, 2003). PAM, on the other hand, focuses on the perceptual ability of learners acquiring non-native speech sound in a foreign setting (Best, 1994b; Best, 1995; Best & Tyler, 2007). In its initial proposal, PAM focused heavily on naïve listeners, but this was later extended to accommodate experienced L2 learners (Best & Tyler, 2007). The L2LP model is specifically aimed at explaining the entire process of L2 acquisition. This model explains the L2 acquisition process from its initial state when learners have to rely on their L1 to the development stage when learners start to establish a separate system for L1 and L2 (Escudero et al., 2014). All of these three perception models hypothesize that the degree of similarity between the L1 and L2 sound contributes to the perception of L2 sounds (Alispahic et al., 2017; Escudero et al., 2014; Escudero & Williams, 2011).

Flege (1995,), however, argued that in order to accurately reproduce an L2 sound, learners would require three competencies. These competencies are, first, the ability to correctly examine the properties of the L2 sound. The second is the ability to differentiate the properties of the sound within L2 and against L1. The third competency is the ability to reproduce the represented L2 sound by learning its motoric gesture required to produce the sounds. The perception process of L2 sounds is said to occur in the first and the second competence. Nevertheless, Flege (1995) points out that SLM is aimed to predict speech perception for advanced learners who have spoken the language for many years and have learned it in a natural setting. As for L2 learners studying English in a non-natural or classroom environment, Rallo Fabra and Romero (2012) suggest that researchers refer to the PAM model developed by Best (1994a). Even though both SLM and PAM have been supported by many studies (Ghaffarvand Mokari & Werner, 2015; Pillai & Salaemae, 2012; Best & Tyler, 2007), Barrios, Jiang, and Idsardi (2016) argue that neither model has so far provided any significant explanation of how the similarity between L1 and L2 is defined.

Vowel inventories of Acehnese

Acehnese has 10 monophthong oral vowels (Durie, 1985; Asyik, 1987; Yusuf, 2013). The vowels are /a/, /i/, /u/, /e/, /o/, /ə/, /ɯ/, /ʌ/, /ɔ/, and /ɛ/ (see Table 1). All Acehnese oral vowels are lax vowels, and thus, differ in terms of length from some English tense vowels. Duration is not a distinctive feature in Acehnese. The Acehnese /ʌ/ and /ɛ/ are identical to English /ʌ/ and /ɛ/ while the Acehnese /i/, /u/ and /ɔ/ are similar to English /i:/, /u:/ and /ɔ:/ but are not lengthened in citation form.

Vowel inventories of Bahasa Indonesia

Bahasa Indonesia (BI), on the other hand, has a smaller vowel inventory compared to Acehnese (see Table 1). BI has six oral vowels /a/, /i/, /u/, /e/, /o/, and /ə/ (Zanten, 1986; Soderberg & Olson, 2008), all of which are lax vowels. All BI vowels exist in Acehnese

since Indonesia has a smaller vowel system. However, there are only two vowels in BI which are similar to English, /i/ and /u/, both of which are shorter than English /iː/ and /uː/.

A Comparison of Acehnese, Indonesian and English Vowels

Standard American English has a total of 10 monophthongs /æ/, /ɪ/, /ʊ/, /ʌ/, /e/, /iː/ /uː/ /ɑː/ /ɔː/ and /ɝ/ (Ladefoged & Johnson, 2014). English vowels differ from Acehnese and BI both in quality and length. Four English vowels are long vowels, /iː/ /uː/ /ɑː/ /ɔː/ while the rest are shorter. Table 1 shows the vowel inventories of Acehnese, BI and American English.

Table 1. Vowel chart comparison across Acehnese (Asyik, 1987), Indonesian (Zanten, 1986) and American English (Ladefoged & Johnson, 2014)

English	Acehnese	Bahasa Indonesia
/iː/	/i/	/i/
/ɪ/	/e/	/e/
/e/	/ɛ/	/ə/
/æ/	/ɯ/	/a/
/ʌ/	/ə/	/u/
/ɑː/	/ʌ/	/o/
/ɔː/	/a/	
/ʊ/	/u/	
/ɝ/	/o/	
/uː/	/ɔ/	

Escudero (2005) suggested that the L1/L2 comparison of vowel inventories could predict the ease of the acquisition of the L2 vowels. Similarly, Best and Tyler (2007) also argued that listeners' relative ease and difficulty in perceiving non-native sounds could be predicted by comparing the phonetic similarities between L1 and L2. In determining similarity, Flege (1987 p. 48) coined the terms' identical', 'similar', and 'new' to classify the L1/L2 phonemic comparisons. The terms can be classified by looking at three accounts (1) the IPA symbol L1 and L2 used to represent its sounds, (2) difference in acoustic features of sounds in both languages, (3) difference in the perception of the two languages. 'Identical' in Flege's term is then defined as sounds that have the same IPA symbols in both languages and have comparable acoustic features while 'similar' is a term referred to sounds with the same IPA symbol in the two languages yet have different acoustic features.

By comparing phonetic similarity across three languages, we could classify which English vowels are considered as similar, new or identical phonemes for Acehnese-Indonesia bilinguals. The comparison and classification can be observed in Table 2. Only two English vowel pairs contain a similar vowel to Indonesian, while there are two vowel pairs containing an identical vowel and three vowel pairs containing a similar vowel to Acehnese. The English /iː/ and /uː/ have comparable sounds in both Acehnese and Indonesian and only differs in term of length. The English vowels /ʌ/ and /ɔː/ have equal sounds in Acehnese, but the former is identical while the letter is similar in terms of length. The English vowel /ɪ/, /ʊ/, /æ/, and /ɑː/ do not have comparable sounds in Indonesian (Karlina et al., 2020) and Acehnese and are classified as 'new'. We then organized the vowel pairs into four categories NNB (New-new to Acehnese-Indonesian), SNA (Similar-new to Acehnese), SNB (Similar-new to Acehnese-Indonesian), and INA (Identical-new to Acehnese). The classification can be observed in Table 3.

Table 2. Comparison of English vowel pairs (VP) to Acehnese and Indonesian based on the framework by Flege (1987)

	Identical-New	Similar-New	New-New
Indonesian	X	i: - ɪ (VP2) u: - ʊ (VP3)	ɛ – æ (VP1) ʌ - ɑ: (VP4) ɔ: - ɑ: (VP5)
Acehnese	ʌ - ɑ: (VP4) ɛ – æ (VP1)	i: - ɪ (VP2) u: - ʊ (VP3) ɔ: - ɑ: (VP5)	X

Table 3. Classification of English vowel pairs by NNB (New-new to Acehnese-Indonesian), SNA (Similar-new to Acehnese), SNB (Similar-new to Acehnese-Indonesian), and INA (Identical-new to Acehnese) categories

Category	Vowel Pairs
INA	ʌ - ɑ: (VP4) ɛ – æ (VP1)
SNB	i: - ɪ (VP2) u: - ʊ (VP3)
SNA	i: - ɪ (VP2) u: - ʊ (VP3) ɔ: - ɑ: (VP5)
NNB	ɛ – æ (VP1) ʌ - ɑ: (VP4) ɔ: - ɑ: (VP5)

METHOD

Respondents

American English Speakers

A 26 year old American speaker provided the stimuli for our recording to be used for the perception task. The speaker worked as an ESL teacher at the Language Center of Syiah Kuala University for one and half years on a teacher exchange program. He had been in Aceh for one year at the time of the recording and considered himself to be monolingual. He was born and raised in Oregon and went to college in Ohio. He described himself as having a Western American accent. Before distributing the perception task to the target participants, the American speaker was recorded reading a list of words containing the target vowel in a CVC context within a carrier sentence of "*I say (target word) again*." The speaker was recorded with a Zoom H5 recorder in a sound-attenuated room at the Language Centre of Syiah Kuala University. The recordings were sampled at 44.000 kHz.

Acehnese-Indonesian Bilingual Learners of English

The participants who took part in the current perception study were 31 second and third-year high school students in Banda Aceh. They were aged between 16 and 18 years

old with an average age of 17 years old. The students comprised 15 females and 16 males. All participants do not have any speech and hearing problems. The high school in question is an Islamic Boarding High School in which students stay in school dormitories and are only permitted to leave the school every two weeks on Sunday. At the school, students are required to speak English to their friends, teachers, and dormitory supervisors, failing which they will be given punishments ranging from having to memorize vocabulary to doing community services. In their first year, students are given three months to code-switch between English and BI before having to use English daily. The students are in the second and third year of their high school and should have accumulated adequate peer exposure to English for about one to two and half years during their time at the school. English is compulsory in high school in Indonesia and is taught three hours per week in the first year and four hours per week in the second and third year. Thus, the second-year students should have learned English at the school for 200 hours while the third-year students are about 360 hours.

In order to ensure their proficiency level, students were selected based on their English examination scores in the final two semesters. Only students with an English score of more than 80 participated in the study. Prior to selection, we distributed questionnaires to confirm that our students met the criteria of bilinguals in Indonesian and Acehnese. Since Acehnese is reported to have a wide variety of dialects (Asyik, 1987), only selected students from one regional dialect were selected. This was done to control the vowel inventory, which might be slightly different across Acehnese dialects. Since most students at the school are from the western part of Aceh, only those with the west Acehnese dialect participated in the study. All students reported using both Acehnese and BI extensively at home, their previous schools, or surrounding neighborhood back in their hometowns. We also made sure that the students speak only Acehnese as their local language because people along the west coast of Aceh also speak Jamee as reported by Yusuf et al. (2021).

Instruments

The target vowels were five English vowel pairs /ɪ/ - /iː/, /æ/ - /e/, /ʌ/ - /ɑː/, /ʊ/ - /uː/, and /ɑː/ - /ɔː/. Whenever appropriate, each of the pairs in this study will be referred to as VP1, VP2, VP3, VP4, and VP5, respectively. Each vowel was recorded in a CVC context word read by an American speaker of English (see Table 4). Each vowel was then extracted and sequenced into targeted pairs in a continuum with an interstimulus interval (ISI) of 700ms and inter-trial interval (ITI) of 2000ms. Højen and Flege (2006) suggested that if the ISI between discriminated items is too short, listeners might perceive the items inaccurately and incomplete. This may lower their score for accuracy.

The experiment was conducted in the AX test format. In this test, the vowel pairs were arranged in four possible stimuli, AA, AB, BB, and BA. A is assigned one vowel such as /ɪ/, and B is another vowel, such as /iː/, within the desired pairs. Each stimulus is then repeated five times, making the total 20 stimuli for each vowel pair and 100 stimuli for all vowel pairs tested. The 20 stimuli of each vowel pair in each task type were randomized and divided into five blocks based on vowel pairs. So VP1 is in the first block with 20 randomized stimuli followed by VP2, VP3, VP4, and VP5. The example of randomized stimuli can be observed in Table 5.

Table 4. English vowel pairs and words used in the study.

	Vowel Pairs	Word Pairs
VP1	**ɪ - i**:	bit - beat
VP2	**æ -** ɛ	bag - beg
VP3	**ʊ - u**:	foot - food
VP4	**ʌ - ɑ**:	bud - bard
VP5	**ɑ**: **- ɔ**:	pot – port

Table 5. AX randomized example for /ɪ/ - /i:/ block

Vowel Pair	Order	Stimuli	Repetition	Total
ɪ - i: **(VP1)**	AA	beat – beat	five rep.	20
	AB	beat - bit	five rep.	
	BA	bit - beat	five rep.	
	BB	bit - bit	five rep.	

Procedures

The perception data were collected by asking the students to listen to instances of the vowel pair stimuli sequenced in the AX format. The test was conducted in a quiet multimedia computer room at the school. Prior to the test, consent forms were distributed, and the purpose of the test was explained to the participants. Prior to this, permission to carry out the study was granted by the school headmaster and head of dormitory supervisors. A permit letter was also obtained from the regional office of the Ministry of Education, Indonesia, to carry out the perception test at the school. In the AX test, students were given papers to mark in the column as 'same' or 'different' based on the stimuli they heard. The first stimulus block was played to familiarize students with the test, after which the students continued the rest of the block until the end of the test.

Data analysis

The answer sheets were then tabulated into an Excel sheet. A score of '1' was given for the correct answer and '0' for the incorrect answer. The score was given based on the stimuli given. If students answered 'same' for the stimuli containing the same vowel pairs, they would get '1' for that stimulus, and if they answered otherwise, they would get '0'. The mean percentage score of each vowel pair was calculated to determine their perception accuracy. To answer research question 1 (RQ1), we compared the mean perception of accuracy across the five vowel pairs using the one-way ANOVA test. An additional post hoc Tukey HSD test was conducted to see which vowel pairs are significant against another. For research question 2 (RQ2), we classified the results from research question one into identical, similar, and new categories and interpreted the data accordingly.

FINDINGS

To what extent do Acehnese-Indonesian bilinguals perceive vowel contrast in English?

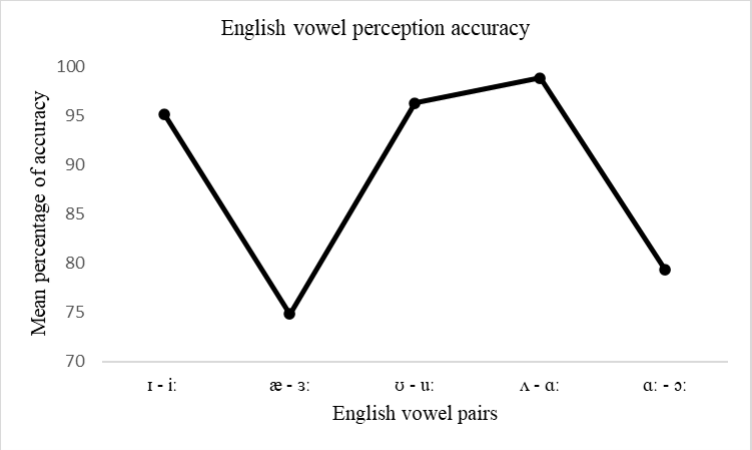

Figure 1. Mean percentage of each vowel pairs in AX Test

Figure 1 shows the mean percentage of perception accuracy across five English vowel pairs /ɪ/ - /iː/, /æ/ - /ɜː/, /ʊ/ - /uː/, /ʌ/ - /ɑː/, and /ɑː/ - /ɔː/ by Acehnese-Indonesian bilingual learners of English. The lowest percentage of accuracy was obtained for /æ/ - /ɜː/ while the highest was observed in /ʌ/ - /ɑː/. Two vowel pairs, /æ/ - /ɜː/ and /ɑː/ - /ɔː/ sit at the bottom of the curve below 80 while the rest vowel pairs are above 95. The mean difference in percentage of accuracy between /æ/ - /ɜː/ and /ɑː/ - /ɔː/ is about 5 points at 75 and 80 respectively. While the mean accuracy for /ʌ/ - /ɑː/ almost reached a perfect score at 99, the values for vowel pairs /ʊ/ - /uː/ and /ɪ/ - /iː/ are slightly lower at 95 and 96 respectively. The mean difference between the vowel pairs with lower accuracy and higher accuracy seems to be quite high at 15 points. However, in order to confirm if the mean difference across vowel pairs is significant, we conducted a one-way ANOVA test to confirm our assumption. The result of the one-way test showed that there is a statistically significant difference in the perception accuracy across the five English vowel pairs tested ($p = 0.00 < 0.05$). A Tukey HSD post hoc test (at a confidence level of 95%) was further conducted to see which of the specified vowel pairs differed from each other. The statistical analysis shows that the vowel pairs /æ/ - /ɜː/ and /ɑː/ - /ɔː/ are significantly different from the vowel pairs /ɪ/ - /iː/, /ʊ/ - /uː/ and /ɪ/ - /iː/ and vice versa. There is no difference in perception accuracy between vowel pairs /æ/ - /ɜː/ and /ɑː/ - /ɔː/ (VP2 × VP5, $p = 0.182$) and no perception difference is observed between the vowel pairs /ɪ/ - /iː/, /ʊ/ - /uː/ and /ɪ/ - /iː/ (VP1 × VP3 $p = 0.982$, VP1 × VP4 $p = 0.369$, VP3 × VP4 $p = 0.715$). It means that the vowel pairs /æ/ - /ɜː/ and /ɑː/ - /ɔː/are in the same subset homogenous group while the vowel pairs /ɪ/ - /iː/, /ʊ/ - /uː/ and /ɪ/ - /iː/are in another subset homogenous group. Our data show that the students' perception is priming to the 'new' vowel pairs (VP2) that are non-existent in both Acehnese and Indonesian and the pairs (VP5) in which one vowel exists only in Acehnese. The perception accuracy was higher to the vowel pairs (VP1 & VP3) in which one of them similar to both Acehnese and Indonesian /iː/ and /uː/ and highest for the vowel pairs with one vowel identical only to Acehnese (VP4).

Table 6. Mean score and Standard Deviation for each English Vowel Pairs

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

To what extent do their existing vowel systems contribute to their perception of English vowel contrast?

In order to see how the degree of similarity between L1/L2 contributes to perception accuracy of English vowel pairs, we plotted the data from Figure 1 into similarity categories we presented in Table 3. Figure 3 clearly shows that the students scored the lowest accuracy for the English vowel pairs that do not exist in Acehnese and Indonesian (NNB) and is followed by the English vowel pairs in which one vowel similar to Acehnese while the other new to Acehnese and Indonesia (SNA). The vowel pairs in which one vowel identical to Acehnese and the other new to Acehnese and Indonesia (INA) surprisingly reached the highest accuracy at 99. The students seem to consistently perceive English vowel pairs in which one of the pairs exists in Acehnese and Indonesian at a similar level, 95 and 96, respectively. A Tukey HSD post hoc (at a confidence level of 95%) shows that the perception accuracy of NNB and SNA vowel pairs are in the same category ($p = 0.182$) while the SNB and INA pairs in another category ($p = 0.369$). It means that the students perform equally between NNB and SNA vowel pairs and SNB and INA pairs.

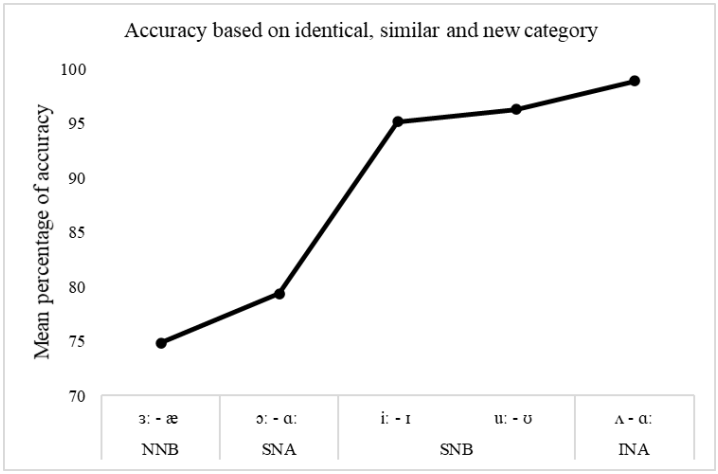

Figure 2. Mean percentage of accuracy-based NNB (New-new to Acehnese-Bahasa Indonesia), SNA (Similar-new to Acehnese), SNB (Similar-new to Acehnese-Bahasa Indonesia), and INA (Identical-new to Acehnese) categories.

DISCUSSION

In RQ1, we seek to learn how Acehnese-Indonesian bilingual learners of English perceive five English vowel contrasts. We specifically look at which vowel pairs are easier

and difficult to perceive. We found that Acehnese-Indonesian bilingual's perception of the English vowel contrasts are unequal and vary for each vowel pair. All of them are able to perceive the English vowel contrast above chance level beyond 50% (Barrios et al., 2016) and 70% (Jia et al., 2006). The students found that some English vowel pairs are difficult to perceive while some others are easier. Our listeners had lower perception accuracy for the English vowel pairs /æ/ - /ɜː/ and /ɑː/ - /ɔː/ and higher perception accuracy for the vowel pairs /ɪ/ - /iː/, /ʊ/ - /uː/, and /ʌ/ - /ɑː/. Lower perception accuracy for the vowel /æ/ and /ɜː/ was also observed in Javanese and Sundanese speakers of English in the previous study by Perwitasari (2018). Our study confirmed that Indonesian speakers continue to experience difficulties in perceiving /æ/ - /ɜː/ contrasts despite different first local language between Acehnese, Javanese and Sundanese Indonesian. However, Perwitasari (2018) also found that Javanese and Sundanese speakers showed higher error rates when perceiving vowels /ʌ/, /ɑː/, and /ɪ/ which is not the case in Acehnese speakers. Javanese speakers even exhibit difficulties on another English vowel /ʊ/. The difference in the methodology we used in the current study may have contributed to this different result. In the current study, we use the AX test, which has been common in perception tests (Barrios et al., 2016; Mora, 2005) while Perwitasari (2018) used a mouse-tracking study. The fact that Acehnese speakers have larger phonemic inventory compared to Javanese and Sundanese may also have contributed to the higher accuracy in most vowel contrast. The study by Iverson and Evans (2007) also confirmed our findings. In their study, Norwegian and German listeners (with larger L1 phonemic inventory) outperformed Italian and Spanish (with smaller L1 phonemic inventory) when identifying English vowels. Even though the current findings seem to negate the findings found by Alispahic et al. (2017) in which listeners with larger phonemic inventories failed to utilize their native vowels in discriminating novel vowel contrast, Alispahic et al. (2017) concluded that larger vowel inventory is not the sole predictor for the acquisition of novel vowel contrast. How close the vowel properties in listeners L1 to the target L2 plays a far more crucial role in determining the ease of discriminating vowel contrast. Acehnese and Indonesian bilinguals have not only large phonemic inventories but also have vowels that are identical to English /ʌ/ and similar to English /ɔː/ and are absent in Indonesian vowel system.

The higher error rates for the vowel contrast /ɑː/ - /ɔː/ in the current study seem to confirm the previous finding in the study by Fata et al. (2017) study. The Acehnese learners of English in their study also produced lack of distinct quality in the vowel pairs /ɑː/ - /ɔː/. However, this conclusion needs to be treated cautiously since the Acehnese speakers in their study also produced /ɪ/ - /iː/ and /ʊ/ - /uː/ contrast similarly but not for the /æ/ - /ɜː/ contrasts. The different proficiency level of the students in our study and that of Fata et al. (2017) may explain the discrepancies in the current findings. However, it is important to highlight that their study focused on production while our study concentrated on perception. While the proficiency level of our listeners is at an intermediate level, their study did not mention the proficiency level of their respondents. Thus, we could not assure for sure as to why inconsistencies happen. However, Flege, Takagi, and Mann (1995) suggested that the state of phonemic perception is not constant but improved over the years as learners receive more input. They found that Japanese speakers living in the United States were eventually able to discriminate between the sound /r/ and /l/ after living in the United States for two years. Learners also learn to pick up more cues to distinguish non-native vowel contrasts as they are exposed to more input from the target language. The respondents in the current study may have been at a certain level of their developmental stage of English vowel contrast acquisition and are at a different stage of language development from the Acehnese-Indonesian in Fata et al. (2017). RQ2, we aim to find out how the similarity or dissimilarity

between the existing Acehnese-Indonesian vowels and English vowels contribute to their perception accuracy. Looking into detail into the similarity between native vowels (Acehnese and Indonesian) and non-native vowels (English), we could determine how the Acehnese-Indonesian bilinguals perceive each non-native vowel contrasts based on their proximity properties to native vowels (see Table 2 and Figure 2). The lowest percentage of perception accuracy were recorded for the pairs /æ/ - /ɜː/ and /ɑː/ - /ɔː/. The English /æ/ - /ɜː/ vowel contrasts are both new to Acehnese and Indonesian bilinguals and do not exist in their lexical inventories. The /ɑː/ - /ɔː/ vowel contrast is not totally new to Acehnese since the English tense /ɔː/ is similar to Acehnese lax /ɔ/. While we predicted that /æ/ - /ɜː/ vowel contrast is difficult to be discriminated since they are new to Acehnese and Indonesian vowels, we did not expect the /ɑː/ - /ɔː/ vowel contrast to exhibit lower scores. The difficulty in perceiving the /æ/ - /ɜː/ vowel contrast can be explained based on two possible factors. First, it can be explained from the L2LP model (Escudero, 2005) which states that L2 learners filter L2 sound through their L1 and whenever available, they will use any closest L1 sound when perceiving non-native sound. Both English /æ/ and /ɜː/ vowels are novel to Acehnese and Indonesian vowel system. They do not have any closest L1 sound to compare the two-novel sounds as an instance of similar or new to their vowel system. Thus, the two sounds are mapped into a single category and perceived as an instance of the same sound. Second, it can also be described based on the optimal perception hypothesis (Escudero, 2005). The Acehnese-Indonesian bilinguals in our study may not have received enough optimal input to pick up the minimal difference properties between the two non-native sounds. During their high school study, they live in the dormitory, and the inputs they get are from their friends who are also learning English or from their English teachers in the class.

The explanation for low perception accuracy for /ɑː/ - /ɔː/ can be best explained from SLM hypothesis #2 by Flege et al. (1995, p. 239) which states that a new category cannot be established if the difference between L2 sound and the closest L1 sound is not well discerned. The Acehnese-Indonesian bilinguals in our study are unable to discern the difference between the new non-native sound /ɑː/ and the closest existing native sound /ɔ/ which result in their poor perception between the new sound /ɑː/ and similar sound /ɔː/. PAM (Best, 1994b) and L2LP (Escudero, 2005) also predict that when non-native vowel contrast is mapped into the single native vowel, the vowels would be difficult to discriminate. This scenario is called the new scenario in L2LP or single category in PAM. The new non-native English /ɑː/ and similar non-native English /ɔː/ may have been mapped into the native Acehnese lax vowel /ɔ/. No statistically significant difference found between the perception of new-new sound /æ/ - /ɜː/ and new-similar sound /ɑː/ - /ɔː/ indicates that Acehnese-Indonesian bilinguals are not only have difficulty to perceive totally new vowel contrast but also vowel contrasts in which one of the pairs contains sound similar only to Acehnese. An in-depth study for this particular vowel pairs on other Acehnese-Indonesian bilinguals needs to be done to confirm this assumption.

The next three vowel contrasts /ɪ/ - /iː/, /ʊ/ - /uː/, and /ʌ/ - /ɑː/ differ from the previous two pairs /æ/ - /ɜː/ and /ɑː/ - /ɔː/. Both English /iː/ and /uː/ have comparable sounds in both Acehnese and Indonesian /i/ and /u/. While the English /iː/ and /uː/ are tense vowels with longer duration, Acehnese-Indonesian /i/ and /u/ are lax vowels with shorter duration. Acehnese-Indonesian bilinguals perceive the two pairs comparably well. They may have successfully used the cue from the target language in terms of duration coupled with reuse of their existing vowel to discriminate between the two sounds. On the other hand, the English ʌ/ has an identical sound in Acehnese /ʌ/ and was found to be the easiest to perceive, reaching almost perfect level. However, statistical analysis shows that the three-vowel

contrasts are in the same homogenous subset and are perceived comparably well by Acehnese-Indonesian bilinguals. For these particular English vowel contrasts, the current listeners have reached optimal listeners and can be said at the end state of their vowel acquisition (Escudero, 2005).

CONCLUSION

In the first question, we address how Acehnese-Indonesian bilinguals perceive five English vowel contrasts. We found that their perception varies across the five pairs tested and reached beyond the chance level threshold of 70% perception accuracy. Both /æ/ - /ɜː/ and /ɑː/ - /ɔː/ vowel contrast are poorly perceived compared to the other three vowel contrasts. Since no statistically significant difference is found between the two pairs, they are in the same category of difficulty. On the other hand, Acehnese-Indonesian bilinguals found /ʌ/ - /ɑː/, /ɪ/ - /iː/, and /ʊ/ - /uː/ are easier to perceive. For the /ʌ/ - /ɑː/ vowel contrast specifically, they reached a near-perfect perception accuracy indicating that they are at the end state of vowel acquisition for this vowel pairs.

In the second question, we specifically look at the perception accuracy of each English vowel contrasts based on their similarity to Acehnese and Indonesian vowels. Perception accuracy of Acehnese-Indonesian bilinguals on English vowel contrast is, to a certain extent, linked with the degree of similarity between native vowels and non-native vowels. The vowel contrasts (/æ/ - /ɜː/) in which both vowels are new to Acehnese-Indonesian vowel system were difficult to discriminate while the vowel contrast (/ʊ/ - /uː/ and /ɪ/ - /iː/) which contain one vowel similar to Acehnese and Indonesian are perceived considerably well. The English vowel contrast (/ʌ/ - /ɑː/) which contains one vowel identical to Acehnese but not to Indonesian was easily perceived. However statistical analysis show that the /ʌ/ - /ɑː/ pairs are in the same level of perception accuracy to /ʊ/ - /uː/ and /ɪ/ - /iː/ pairs. However, the Acehnese-Indonesian bilinguals had difficulties to discriminate the vowel pairs (/ɑː/ - /ɔː/) which contains one vowel only similar to Acehnese but not to Indonesian. It seems that they fail to optimally reuse the knowledge of their native vowel to perceive non-native English vowel contrast. In fact, their difficulty is in the homogenous subset of the novel non-native vowel pairs (/æ/ - /ɜː/).

The current findings contribute to our understanding of the perception of English vowel contrast by Acehnese-Indonesian bilingual learners of English. It sheds light on which vowel contrasts are easier to discriminate and which are difficult. It also provides a glance at how the degree of similarity between Acehnese-Indonesian and English vowels could predict the relative degree of perception accuracy of non-native English vowel contrast. The findings suggest that Indonesian Islamic boarding high schools need to continue providing optimal input to the students during their study at the school so that they could learn to pick up the cues and acoustic property difference of the English vowel absent in their native vowels.

REFERENCES

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

Asyik, A. G. (1987). *A Contextual Grammar of Acehnese Sentences.* Michigan: University of Michigan.

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

Best, C. T. (1994a). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman; & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T. (1994b). Learning to perceive the sound pattern of English. In Lipsitt, L. P., & Rovee-Collier, Carolyn K. (Eds.), *Advances in infancy research*, Vol. 9 (pp. 217-304). New Jersey: ABLEX Publishing Corporation.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-206). Baltimore: York Press.

Best, C. T., & Tyler, M. D. (2007). Non-native and second-language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins Publishing.

Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and Ultimate Attainment in the Pronunciation of a Foreign Language. *Studies in Second Language Acquisition, 19*(4), 447-465. doi:10.1017/S0272263197004026

Elvin, J., Williams, D., & Escudero, P. (2016). Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *The Journal of the Acoustical Society of America, 140*, 576-581. doi: 10.1121/1.4952387

Escudero, P. (2005). Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization. (Doctoral dissertation, Utrecht University & LOT). https://dspace.library.uu.nl/handle/1874/7349

Escudero, P., Sisinni, B., & Grimaldi, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *The Journal of the Acoustical Society of America, 135*, 1577-1584. doi: 10.1121/1.4864477

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

Evans, B. G., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics, 68*, 15-31.

Fabra, L. R. (2005). Predicting ease of acquisition of L2 speech sounds. A perceived dissimilarity test. *Vigo International Journal of Applied Linguistics* (2), 75-92.

Fata, I. A., Fitrian, F., Mohammad, T., & Yusuf, Y. Q. (2017). *Acoustic Analysis on English Oral Vowels Produced by Acehnese Speakers from Aceh Besar by Using PRAAT Software.* Paper presented at the 7th Annual International Conference (AIC), Banda Aceh.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47-65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). Baltimore: York Press.

Flege, J. E. (1999). The relation between L2 production and perception. *14th International Congress of Phonetic Sciences, 2*, 1273-1276.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. S. Meyer & N. O. Schiller (Eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities* (pp. 319-355). Berlin ;Boston: De Gruyter Mouton.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese Adults can Learn to Produce English /I/ and /l/ Accurately. *Language and Speech, 38*, 25-55. doi: 10.1177/002383099503800102

Mokari, P. G., & Werner, S. (2017). Perceptual assimilation predicts acquisition of foreign language sounds: The case of Azerbaijani learners' production and perception of Standard Southern British English vowels. *Lingua, 185*, 81-95.

Højen, A., & Flege, J. E. (2006). Early learners' discrimination of second-language vowels. *The Journal of the Acoustical Society of America, 119*(5), 3072-3084.

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

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

Karlina, Y., Rahman, A., & Chowdhury, R. (2020). Designing Phonetic Alphabet for Bahasa Indonesia (PABI) for the teaching of intelligible English pronunciation in Indonesia. *Indonesian Journal of Applied Linguistics, 9*(3), 724-732.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606-608.

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics* (7th Edition ed.). Cengage Learning.

Larson-Hall, J. (2006). What does more time buy you? Another look at the effects of long-term residence on production accuracy of English vertical bar r vertical bar and vertical bar 1 vertical bar by Japanese speakers. *Language and Speech, 49*, 521-548. doi: 10.1177/00238309060490040401

Mack, M. (1988). Sentence processing by non-native speakers of English: Evidence from the perception of natural and computer-generated anomalous L2 sentences. *Journal of Neurolinguistics, 3*(2), 293-316.

Mora, J. C. (2005). *Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and nonwords by Spanish/Catalan bilingual learners of English.* Paper presented at the ISCA Workshop on Plasticity in Speech Perception.

Mora, J. C., & Nadeu, M. (2012). L2 effects on the perception and production of a native vowel contrast in early bilinguals. *International Journal of Bilingualism, 16*(4), 484-500.

Peperkamp, S., & Bouchon, C. (2011). *The relation between perception and production in L2 phonological processing.* Paper presented at the Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH.

Perwitasari, A. (2013). Slips of the ears: Study on vowel perception in Indonesian learners of English. *Jurnal Humaniora, 25*(1), 103-110.

Perwitasari, A. (2018). Perception of English vowels by Javanese and Sundanese speakers: A mouse-tracking study. *Wacana, 19*(1), 219-234.

Pillai, S., & Salaemae, W. (2012). An instrumental analysis of English monophthongs produced by Thai EFL learners. *Pertanika Journal of Social Science and Humanities, 20*, 1143-1160.

Rallo Fabra, L., & Romero, J. (2012). Native Catalan learners' perception and production of English vowels. *Journal of Phonetics, 40*, 491-508. doi: 10.1016/j.wocn.2012.01.001

Roach, P. (2009). *English phonetics and phonology: a practical course* (Vol. 4th). Cambridge: New York: Cambridge University Press.

Thomson, R. I. (2012). Improving L2 Listeners' Perception of English Vowels: A Computer-Mediated Approach. *Language Learning, 62*, 1231-1258. doi: 10.1111/j.1467-9922.2012.00724.x

Yamada, R.A (1995). Age and acquisition of second language speech sounds perception of American English /r/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience*, (pp. 305–320). Baltimore: York Press

Yusuf, Y. Q., Fata, I. A., & Karwinda, S. (2021). Oral monophthong vowel qualities of the Jamee language in Aceh. *Indonesian Journal of Applied Linguistics*, *10*(3).

Zanten, E. V. (1986). Allophonic Variation in The Production of Indonesian Vowels. *Perception, 142*, 427-446.