

PRODUCTION OF VOWEL /ə/ IN SUNDANESE BY JAPANESE NATIVE SPEAKERS IN FIRST EXPOSURE

Rike Febriyanti^{1*}; Lailatul Husna²

^{1,2}Faculty of Cultural Studies, Japanese Language Education, Brawijaya University
Jl. Veteran, Ketawanggede, Malang, Jawa Timur 65145, Indonesia
¹febriyanti_rike@ub.ac.id; ²husnalaila238@ub.ac.id

Received: 26th October 2020/Revised: 15th December 2020/Accepted: 17th December 2020

How to Cite: Febriyanti, R. & Husna, L. (2020). Production of vowel /ə/ in Sundanese by Japanese native speakers in first exposure. *Lingua Cultura*, 14(2), 225-231. <https://doi.org/10.21512/lc.v14i2.6763>

ABSTRACT

The research described native Japanese speakers' perception of Sundanese vowel /ə/ after the first exposure to a controlled naturalistic input of conversation. The research worked in respect of Brown's model of L2 speech perception and L1 feature geometry, which sought to relate theories of segmental phonology to L2 speech perception and the first exposure treatment. Some Sundanese native speakers conducted a conversation that contained the /ə/ vowel in front of five Japanese native speakers with no prior exposure to Sundanese. Therefore, the researchers had collected speech data from five L1 Japanese native speakers (three males, two females, Mage = 22, SD = 2,1). The Japanese were asked to listen to the short conversation and imitate vowel /ə/, which did not exist in the Japanese language vowel inventory. The observation confirmed Brown's hypothesis that L2 perception of /ə/ vowel was constrained by the L1 feature geometry. L1 Japanese language phonological properties worked as a perceptual filter to Sundanese L2 input, causing the Japanese L2 learners to perceive only the vowel discriminated by phonological features presented in Sundanese. The data show that the Japanese native speakers are able to overcome the perceptual filters so they can produce various frequencies of vowel /ə/, which are statistically similar to the frequency produced by Sundanese native speakers. The research implies that the possibility of learning new sounds from an entirely new language is wide open when the learner is able to pass through the L1 perceptual filter.

Keywords: Sundanese, Japanese language, native speaker, first exposure, second language acquisition

INTRODUCTION

In the field of second language acquisition (SLA), second language (L2) adult learners face many difficulties when working very hard to learn L2 non-native sounds (Steele, 2009; Pfenninger & Singleton, 2019). This challenge is noticeable in the adult L2 learners' language creation of what is known as accented speech (Alwohaibi, 2019). Older children and adults show language-specific patterns, as noted by Curtin and Werker (2018). Moreover, adult L2 learners work very hard to achieve an L2 native-like performance and usually speak with a foreign-accented L2 even when they live for years in an L2 environment (Alwohaibi, 2019). Steele (2009) has stated that the infant and the adult could comprehend the same speech in the almost similar way, nor could the L2 learner or bilingual comprehend L2 or L1 speech in accurately the same path as native monolinguals of both languages.

Therefore, adults' L1 language-specific experience hinders the observation of L2 speech contrasts that are phonologically dissimilar from those of the listener's native language. This imprecise perception of L2 non-native sounds can cause adult L2 learners to struggle with specific L2 sound contrasts (Melnik, 2019).

The experience of foreign speech (L2) sounds by L2 learners is highly influenced by the following elements: the starting age of L2 acquisition (Alwohaibi, 2019), the total of meaningful exposure to the L2 (de Leeuw, 2019), and the L1 vowel and consonant system (Elvin, Escudero, & Vasiliev, 2014). Preceding research has shown that L2 learners who have a lesser number of L1 vowels suffer many problems in perceiving an L2 with a larger number of vowels (Elvin, Escudero, & Vasiliev, 2014). This phenomenon is proved when L2 learners who do not possess the sound contrasts in their first language. Souza et al. (2017) have analyzed L1 and L2 vowel systems of native German, Spanish,

Mandarin, and Korean speakers. The result shows that the form of the L1 vowel system and its perceived relation to vowels influence the L2 vowel production and perception. The result confirms that, regardless of the differences in their vowel systems, learners' perceptions are precisely calculated by specified acoustic comparison between L1 and L2 sounds. An example of L2 sound perception is the Speech Learning Model (SLM).

Models of L2 sound perception, such as the Speech Learning Model (SLM) and the Second Language Linguistic Perception (L2LP) (Alwohaibi, 2019) and the Second Language Linguistic Perception (L2LP) model (van Leussen & Escudero, 2015), have suggested that acoustic matches between L1 and L2 sounds is important in cross-language speech perception. However, the exact predictions for L2 perception between the two previous models are different. Generally, the SLM tries to describe that second language learning's success is mostly affected by the perceived phonetic similarities between the L1 and L2 sounds. Inside the SLM, the noticed phonetic similarities are divided into identical, similar, and new sounds, as noted by (Carlet & Souza, 2018). An identical sound is distinguished by similar International Phonetic Alphabet (IPA) symbols. It has no essential acoustic dissimilarities between the L1 and L2 sounds. An L2 sound is defined as a similar sound if characterized by the same IPA symbol as a sound in the L1 and if the difference is in the diacritics only. A new sound is defined as an L2 sound, which is not used in the L1 differs auditory from the closest L1 sound, and for which the IPA origin symbol is different (Alwohaibi, 2019). The SLM shows that L2 learners will have no important problem making and perceiving a similar sound when they cannot notice the L1 and L2 sound differences. However, L2 learners will be less successful in receiving similar sounds since the similarity between L1 and L2 sounds will block the phonetic category's formation. They will be successful in perceiving new vowels as it motivates them to produce new uncategorizable speech sounds.

Unlike the SLM, the L2LP states a prediction that the L2 learners will confront different kinds of a perceptual struggle depending on how the perception language rules in the L1, which matches the ideal L2 perception. In the L2LP, the awareness of an L1 contrast is divided into three situations: new, sub-set, and similar. A new situation occurs when the L1 perception grammar results in less perceptual classifications than the necessary perception of the L2. As a result, the L2 setting creates phonological variances that do not exist in the L1 (Nimz & Khattab, 2019). For example, Spanish learners of English consider the two English sounds /i/-/i/ onto a specific native sound set /i/. The perception of new sounds is believed to be the most problematic situation. It includes not only the production of new classifications and perceptual mappings but also the combination of the newly classified dimensions into the already classified dimensions (Nimz & Khattab, 2019). The

sub-set scenario occurs if the L1 perception language rule produces more categories than the necessary perception of the L2. Therefore, the L2 categories make a sub-set of L1 categories. For example, Dutch English learners consider the Spanish /i/ into two native classifications /i/ and /i/. In the comparable situation, the L1 perception grammar makes the same quantity of classification as the target of L2 grammar since the L1 and L2 categories are phonologically equivalent. For instance, L1 Canadian English speakers map the Canadian French sounds /ɛ/ and /æ/ onto /ɛ/ and /æ/ in the native categories.

Perwitasari (2018) has noted that Sundanese is used daily by about 34 million individuals in Indonesia, the second most commonly spoken language in Indonesia after Javanese. Sundanese is spoken in the western part of the island of Java (Figure 1). In Sundanese, it can be found at least four main dialects: Banten, Bogor/Karawang, Priangan, and Cirebon (Muslim et al., 2010). The Banten dialect is spoken in several cities around Banten; the Bogor/Karawang dialect is uttered in several big cities such as Tangerang, Bogor, Purwakarta, Krawang, and Subang; the Priangan dialect is spoken in Priangan; and the Cirebon dialect is spoken in Cirebon, Brebes, and Cilacap. Sundanese is distinguished into four speech levels: high *basa lemes*, neutral *basa sedeng*, everyday *basa kasar*, and low *basa tjohag* (Perwitasari, Klamer, & Schiller, 2016). The students who participate in the current research speak the Bogor/Karawang dialect of Sundanese.



Figure 1 Map of Dialects of Sundanese Language in West Java (Muslim et al., 2010)

Sundanese sound inventory has seven vowels. Perwitasari (2018) has mentioned that Sundanese vowels are classified as high front /i/, high central /i/, high back /u/, mid front /e/- /ɛ/, mid central /ə/, mid back /o/-/ɔ/, and central low /a/. Sundanese vowel inventory is illustrated in Figure 2.

Meanwhile, the Japanese vowel system comprises five distinguishable form /i, e, a, o, u/, which form five short (1-mora) and long (2-mora) pairs (Yazawa & Kondo, 2019). Japanese has only five vowels in its vowel inventory; a system is quite common among many natural languages in the world (Sanjaya, 2018). Figure 3 shows the Japanese vowel

inventory.

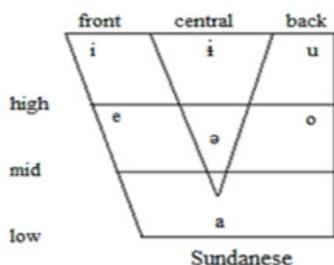


Figure 2 Traditional View of Sundanese Vowel Inventory (Perwitasari, 2018)

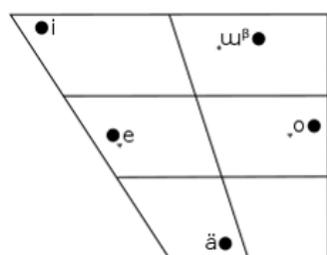


Figure 3 Japanese Vowel Inventory (Yazawa & Kondo, 2019)

The researcher works as a Japanese language lecturer at Brawijaya University and is currently working with native Japanese speakers. There are two kinds of Japanese native speakers at the university. The first is those who are working as Japanese language lecturers, and the second is the Japanese native speakers who are joining the apprenticeship program at Brawijaya. In daily interaction among the native speakers, lecturers, and students, the researcher finds that the Japanese native speakers often casually learn several local languages brought by the students of Japanese language education study program such as Javanese, Sundanese, or Madurese. The Japanese native speakers learn to produce unique vowels and consonants from the local language. The researcher has observed that they struggle to imitate some vowels and consonants that they do not possess in their mother tongue. One incident that sparks the researcher's interest happens when the Japanese native speakers struggle to imitate vowel /ə/ from various Sundanese words. That moment has inspired the researcher to examine how Japanese native speaker's perception of vowel /ə/ of Sundanese in first exposure. The research is the beginning of a series of studies to investigate Japanese native speakers' perception of vowels and consonants in local Indonesian languages. The researcher hopes that the research and her forthcoming studies will help mutual understanding between Japanese native speakers who learn local languages and the students who learn Japanese.

Currently, in the field of Second Language Acquisition, studies of first exposure to L2 contribute to bridging gaps in the literature about the crucial

first minutes/hours of naturalistic L2 contact. So, the studies contribute the knowledge of the adults' early stages of L2 perception, perceiving, and understanding. In turn, they offer ideas for comparison in the middle of L1 child acquisition and L2 adult acquisition. The research aims to investigate adult L2 learners' perceptual abilities of non-native L2 sounds after first exposure under the feature geometry model of Alwohaibi (2019). Therefore, the literature in the research on first exposure studies will present a general view of the major empirical studies of first exposure studies conducted so far, such as Carroll (2007, 2012, 2014), Gullberg et al. (2010), Han and Liu (2013), Rast (2008, 2010) and Park & Han (2008) Alwohaibi (2019).

SLA researchers in the middle of many other scholars have organized L2 research examining the effect of an organized artificial and/or naturalistic input with adult L2 learners to analyze adult L2 learners' input processing abilities. Many studies have also studied first exposure through implicit learning, such as Gullberg, Roberts, and Dimroth (2012), and Cox (2019). They have examined L1 Dutch learners' first exposure to L2 Mandarin with a realistic input of a 14 minutes Mandarin weather report as the treatment. Cox's (2019) data have shown a positive correlation between the frequency of input and results' accuracy. Such data show that adult L2 potential acquisition skills are far higher than might be expected at first exposure. Han and Liu's findings in 2013 contradict Gullberg's findings (Haghani & Maftoon, 2016) in exposed English and Japanese L1 learners to L2 Mandarin through the use of ten video episodes with varied themes such as ordering food in a restaurant and bargaining in a shop. Each episode lasts for three minutes. Han and Liu (2013) have concluded that these English and Japanese L1 learners of Mandarin struggle through all of the input processing tasks, including free recall, comprehension, note-taking, elicited imitation, and a working memory test. Carroll's research in 2012 and 2014 in Wen, Biedroń, and Skehan (2017) is performed by exposing native speakers of English to L2 German to evaluate for their segmentation abilities after first exposure. The treatment contains audio-visual stimuli showed as a name-learning task. The data indicate that learners could segment the L2 speech stream with around 90% accuracy rate and also map sound tokens to referents, even with low contact frequency. It advises that segmentation abilities are evident in L2 learners even after limited numbers of minor exposure. Other scholars have also analyzed segmentation after first exposure, such as Rast in Berthelsen et al. (2020). They also claim that the L2 learner's ability to segment linguistic components from the L2 speech stream is evident and discussed the linguistic variables of input that may affect the L2 learner's segmentation ability, such as frequency of input and gestures (Archibald, 2017).

Onishi (2016) has tested L2 learners' linguistic knowledge of phonotactic constraints after listening to a short auditory record in three treatments and

tests. The purpose of that research is to test whether adult English speakers could obtain phonotactic regularities that do not exist in English. Onishi (2016) has claimed that their findings demonstrate the phonotactic constraints that are rapidly learned from brief auditory experience and that some constraints are more easily learned than others. The other research is by Altenberg in Alwohaibi (2019), who creates three experiments to examine the acquisition of English word-initial consonant clusters by native speakers of Spanish. These experiments consist of a metalinguistic judgment task, a perception task, and a production task. It suggests that beginning, as well as advanced L2 learners, show an accurate knowledge of English phonotactics and that L1 transfer does not play a role in the learners' perception. It should be noted that first exposure studies do contribute to elucidating adult L2 learners' linguistic abilities when faced with an unknown language for the first time. As mentioned earlier, there are a limited number of studies addressing this issue in SLA. Therefore, future research needs to be addressed this matter.

METHODS

The research is based on the hypothesis that L1 Japanese feature geometry will mediate between the incoming acoustic stimuli of the speech stream of Sundanese L2, sorting the stimuli into phonemic perceptual categories (Alwohaibi, 2019). Therefore, the researchers have collected speech data from five L1 Japanese native speakers (three males, two females, Mage = 22, SD = 2,1), although the gender and age of the participants are not variables. The participants speak the Japanese language as their first language. All the participants demonstrate normal speech and hearing abilities. Auditory stimuli have comprised of one vowel /ə/ in two words, 'hareudang' and 'heureuy' of Sundanese. Each stimulus is recorded using Praat at 44,1 kHz. Participants are tested individually in a sound-attenuated room. Before the experiment, the participants have filled in a demographic questionnaire and signed a consent form.

The treatment includes an audiovisual input of an original short conversation in Sundanese. The short conversation is completely in Sundanese and lasted for approximately two minutes. The research participants are told to ensure that their speakers are working appropriately. They sit in a quiet place before watching the short conversation and before proceeding to listen to the test items. The short conversation is delivered by two male native Sundanese. In order to be as realistic as possible, each test word containing the vowel /ə/ under examination is inserted in the following short conversation. If the Japanese native speakers do not differentiate between two sounds, the assumption is that the phonological representations of this learner's L1 lack the necessary existing structure to facilitate differentiating between the two sounds.

The Japanese native speakers, then, are asked

to listen to the short conversation several times. When the Japanese native speakers confirm that they could listen and grasp the vowel /ə/, then the researchers stop the record. The Japanese native speakers are asked to imitate the vowel /ə/ in one shot. Using Praat, the researchers isolate vowel /ə/ produced by both Sundanese and Japanese native speakers. After isolating the vowel, the researchers analyze the nature of vowel /ə/ produced by both groups.

RESULTS AND DISCUSSIONS

First, since vowel /ə/ is a sound that does not exist in the phonemic inventory of the Japanese language, so participants are not expected to distinguish these sounds accurately. Statistical analysis is implemented using the statistical software package SPSS v.20 for Windows. First, the data are described using descriptive statistics (mean, median, and standard deviation), which are analyzed to describe the numerical quantitative variables. In general, there are apparent differences between the Japanese native speaker's vowel /ə/ production and the Sundanese speaker's vowel /ə/ production. Table 1 shows the statistical data of the vowel /ə/ length production by Japanese native speakers and the Sundanese speakers.

Table 1 The Vowel /ə/ Length Production by Japanese Native Speakers and The Sundanese Speakers

No.	Sundanese Native Speaker	No.	Japanese Native Speaker
S1	0,192404	J1	0,218811
S2	0,182454	J2	0,207269
S3	0,193345	J3	0,258146
S4	0,179921	J4	0,325848
S5	0,18549	J5	0,279503

Table 2 The Statistical Analysis of Vowel /ə/ Length Production by Japanese Native Speakers and The Sundanese Speakers

t-Test: Paired Two Sample for Means		
	0,192404	0,218811
Mean	0,1853025	0,2676915
Variance	3,39305E-05	0,002421228
Observations	4	4
Pearson Correlation	-0,250514683	
Hypothesized Mean Difference	0	
df	3	
t Stat	-3,232329591	
P(T<=t) one-tail	0,024065261	

Table 2 The Statistical Analysis of Vowel /ə/ Length Production by Japanese Native Speakers and The Sundanese Speakers (Continued)

t Critical one-tail	2,353363435
P(T<=t) two-tail	0,048130521
t Critical two-tail	3,182446305

From Table 2, it could be understood that the average length of vowel /ə/ produced by Sundanese native speakers is 0,1853025 seconds. On the other hand, the Japanese native speakers take 0,2676915 seconds to produce vowel /ə/ in speaking *hareudang*. The data variances between the two groups are different. The data variance in Sundanese native speakers is very high at 3,39305, while in Japanese native speakers is at 0,002421228 in both four observations with the degree of freedom (df) of 4-1 is 3. In addition, there is a descriptive Pearson correlation, namely -0,250514683, so it can be said that the relationship is not negative and far. Based on these results, it is known that the t stat is -3,232329591. The value obtained is the same as in the paired t-test material. The hypothesis used is the two-way hypothesis, so that it uses two tails. The result t table is 2,093024 with a p-value of 0,048130521. Because the p-value is smaller than alpha 5% or by looking at | t count | > t table, then the decision is Reject H0. Because H0 is rejected, so it is concluded that there is a significant difference between the vowel /ə/ produced by Sundanese and Japanese native speakers.

After observing the two groups in speaking *hareudang*, Table 3 and 4 present their performance in another word *heureuy*.

Table 3 The Vowel /ə/ Length Production by Japanese Native Speakers and The Sundanese Speakers

Vowel length /ə/ in <i>Heureuy</i>			
No.	Sundanese Native Speaker	No.	Japanese Native Speaker
S1	0,098211	J1	0,08474
S2	0,1082	J2	0,246178
S3	0,858211	J3	0,087964
S4	0,08256	J4	0,2637231
S5	0,818293	J5	0,0240437

Table 4 The Statistical Analysis of Vowel /ə/ Length Production by Japanese Native Speakers and The Sundanese Speakers

t-Test: Paired Two Sample for Means			
		0,192404	0,218811
Mean		0,185303	0,267692

Table 4 The Statistical Analysis of Vowel /ə/ Length Production by Japanese Native Speakers and The Sundanese Speakers (Continued)

Variance	3,39E-05	0,002421
Observations	4	4
Pearson Correlation	-0,25051	
Hypothesized Mean Difference	0	
df	3	
t Stat	-3,23233	
P(T<=t) one-tail	0,024065	
t Critical one-tail	2,353363	
P(T<=t) two-tail	0,048131	
t Critical two-tail	3,182446	

From Table 3 and 4, it could be understood that the average length of vowel /ə/ produced by Sundanese native speakers is 0,185303 seconds. On the other hand, the Japanese native speakers take 0,267692 seconds to produce vowel /ə/ in speaking *hareudang*. The data variances between the two groups are different. The data variance in Sundanese native speakers is very high at 3,39E-05, while in Japanese native speakers is at 0,002421 in both four observations with the degree of freedom (df) of 4-1 is 3. In addition, there is a descriptive Pearson correlation, namely -0,25051, so it can be said that the relationship is not negative and far. Based on these results, it is known that the t stat is -3,23233. The value obtained is quite similar to the paired t-test material. The hypothesis used is the two-way hypothesis, so that it uses two tails. The result t-table is 3,182446 with a p-value of 0,048131. Because the p-value is smaller than alpha 5% or by looking at | t count | > t-table, then the decision is Reject H0. Because H0 is rejected, so it is concluded that there is a significant difference between the vowel length /ə/ produced by Sundanese and Japanese native speakers.

Tables 5 and 6 will describe the data of the frequency of vowel /ə/ produced by Sundanese and Japanese native speakers.

Table 5 The Frequency of Vowel /ə/ Production by Japanese Native Speakers and The Sundanese Speakers

Frequency of Vowel /ə/ in <i>Heureuy</i> (Hz)			
No.	Sundanese Native Speaker	No.	Japanese Native Speaker
S1	141	J1	146
S2	139	J2	210,8
S3	144	J3	141,8
S4	142	J4	96
S5	144	J5	124

Table 6 The Statistical Analysis of The Frequency of Vowel /ə/ Production by Japanese Native Speakers and The Sundanese Speakers

t-Test: Paired Two Sample for Means		
	141	146
Mean	142,25	142,75
Variance	5,583333	2354,25
Observations	4	4
Pearson Correlation	-0,70577	
Hypothesized Mean Difference	0	
df	3	
t Stat	-0,01991	
P(T<=t) one-tail	0,492681	
t Critical one-tail	2,353363	
P(T<=t) two-tail	0,985362	
t Critical two-tail	-0,01991	

From Table 5 and 6, it could be understood that the average frequency of vowel /ə/ produced by Sundanese native speakers is 142,25 Hz. On the other hand, the Japanese native speakers have reached 142,75 Hz in producing vowel /ə/ in speaking *hareudang*. The data variances between the two groups are different. The data variance in Sundanese native speakers is very high at 5,583333, while in Japanese native speakers is at 2389,397 in both four observations with the degree of freedom (df) of 4-1 is 3. In addition, there is a descriptive Pearson correlation, namely -0,70577, so it can be said that the relationship is negative and far. Based on these results, it is known that the t stat is -0,01991. The value obtained is quite similar to the paired t-test material. The hypothesis used is the two-way hypothesis, so that it uses two tails. The result t-table is -0,01991 with a p-value of 0,985362. Because the p-value is larger than alpha 5% or by looking at $|t \text{ count}| > t\text{-table}$, then the decision is Accept H_0 . Because H_0 is accepted, so it is concluded that there is no significant difference between the frequency of vowel /ə/ produced by Sundanese and Japanese native speakers.

CONCLUSIONS

In summary, the data show two different facts. First, the present results reveal significant differences between the performances of the Japanese native speakers and the Sundanese native speakers in producing the length of vowel /ə/. The Japanese native speakers produce significantly longer vowel /ə/ than the Sundanese native speakers. In this first exposure study, the researchers predict that this longer time taken by the Japanese native speaker is a compensation for the absence of vowel /ə/ in the Japanese language.

Therefore, the Japanese native speakers need more time to perceive, think, and produce vowel /ə/. Although statistically different, the longer period of time used by Japanese native speakers to produce vowel /ə/ shows the group's potential to produce a new vowel in the first exposure. On the other hand, the data presented that the frequency production of vowel /ə/ by both groups is statistically similar. The Japanese native speakers could position their tongue is halfway between a close vowel (a high vowel) and a mid-vowel, between a front vowel and a back vowel, and their lips are unrounded. They produce the vowel /ə/ in such a way, although they do not have any experience producing the vowel /ə/ before in their daily life.

Thus, the research hypothesis that L1 feature geometry changes L2 perception accuracy of non-native vowel /ə/ directing to imprecise L2 perception of the non-native vowel contrasts is partially supported. The research shows that Japanese native speakers could not produce vowel length /ə/, which is absent in the Japanese feature geometry. It proves Brown's hypothesis that speakers of a given L1 can only perceive those non-native contrasts distinguished by a feature present in their L1 grammar. However, this data display that the Japanese native speakers have the potential to produce a set of new vowel frequency reach statistically similar frequency to the frequency of vowel /ə/ produced by the Sundanese native speakers. This aspect confirms Gullberg's data, which describes a positive correlation between the input frequency and results' accuracy. The data found by Gullberg indicate that adult L2 potential acquisition skills are far higher than might be expected at first exposure.

Based on the current results, L1 Japanese language phonological properties work as a perceptual filter that filters the Sundanese L2 input. It causes the Japanese L2 learners to perceive only the vowel that is discriminated by phonological features in Sundanese. This perceptual filter works prominently at first exposure to L2 Sundanese, which is obvious in the results of the Japanese native speakers with no prior exposure to Sundanese. The data gathered by this research show that the Japanese native speakers are able to overcome the perceptual filters so they can produce various frequencies of vowel /ə/, which are statistically similar to the frequency produced by Sundanese native speakers.

One source of weakness of the research that might have influenced the outcome is the small number of research participants. Therefore, more research participants, either from Japanese language native speakers or Sundanese native speakers, will produce a different and more statistically credible outcome. So it is recommended that upcoming research should take into account including larger sample size. Another weakness of the research is the setting which does not allow the participants' voices to be recorded in a quiet environment. Because of the quiet environment, the researchers could not identify all aspects of the sounds accurately. Furthermore, it is suggested that all recordings be conducted at least in a quiet room

or in a laboratory environment to reduce any factors that might influence the participants' performance. In addition, future researches are suggested to plan experiments, which combine speech production tasks. The research opens a pathway for future research about foreign language learners' potential capacity in learning other vowels or consonants in other ethnic languages in Indonesia.

REFERENCES

- Alwohaibi, H. (2019). *Investigating native English speakers' perception of novel Arabic phonemes after first exposure*. (Master Thesis). Newcastle: Newcastle University.
- Archibald, J. (2017). *Second language processing and linguistic theory*. Oxford: Oxford Research Encyclopedia in Linguistics.
- Berthelsen, S. G., Horne, M., Shtyrov, Y., & Roll, M. (2020). Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence. *Brain Research*, 1729, 1-15. <https://doi.org/10.1016/j.brainres.2019.146614>.
- Carlet, A., & Souza, H. K. (2018). Improving L2 pronunciation inside and outside the classroom. *Ilha Do Desterro A Journal of English Language, Literatures in English, and Cultural Studies*, 71(3), 99-123. <https://doi.org/10.5007/2175-8026.2018v71n3p99>.
- Cox, J. G. (2019). Multilingualism in older age: A research agenda from the cognitive perspective. *Language Teaching*, 52(3), 360-373. <https://doi.org/10.1017/S0261444819000193>.
- Curtin, S., & Werker, J. F. (2018). Primir on Tone. *Frontiers in Psychology*, 9, 1-2. <https://doi.org/10.3389/fpsyg.2018.01007>.
- de Leeuw, E. (2019). Phonetic Attrition. In S. Monika & B. Kopke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 1-32). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.17>.
- Elvin, J., Escudero, P., & Vasiliev, P. (2014). Spanish is better than English for discriminating Portuguese vowels: Acoustic similarity versus vowel inventory. *Frontiers in Psychology*, 5, 1-8. <https://doi.org/10.3389/fpsyg.2014.01188>.
- Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *IRAL - International Review of Applied Linguistics in Language Teaching*, 50(4), 239-276. <https://doi.org/10.1515/iral-2012-0010>.
- Haghani, M., & Maftoon, P. (2016). The impact of learning styles on the Iranian EFL learners' input processing. *Journal of Language and Translation*, 6(2), 11-26.
- Han, Z., & Liu, Z. (2013). Input processing of Chinese by 'ab initio' learners. *Second Language Research*, 29(2), 145-164. <http://doi.org/10.1177/0267658313479359>.
- Melnik, G. A. (2019). *Issues in L2 phonological processing*. Paris: Universite Paris. Retrieved from <https://tel.archives-ouvertes.fr/tel-02304656/document>.
- Muslim, D., Haerani, E., Shibayama, M., & Yamaguchi, H. (2010). Language mapping based on Geomorphology in western part of Java, Indonesia. *Memoirs of Osaka Kyoiku University, Ser.III*, 58(2), 11-18.
- Nimz, K., & Khattab, G. (2019). On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research*, 36(4), 623-652. <https://doi.org/10.1177/0267658319828424>.
- Onishi, H. (2016). The effects of L2 experience on L3 perception. *International Journal of Multilingualism*, 13(4), 459-475. <https://doi.org/10.1080/14790718.2016.1217604>.
- Perwitasari, A. (2018). Perception of English vowels by Javanese and Sundanese speakers: A mouse-tracking study. *Wacana*, 19(1), 219-234. <https://doi.org/10.17510/wacana.v19i1.607>.
- Perwitasari, A., Klamer, M., & Schiller, N. O. (2016). Formant frequencies and vowel space area in Javanese and Sundanese English language learners. *3L: The Southeast Asian Journal of English Language Studies*, 22(3), 141-152. <https://doi.org/10.17576/3L-2016-2203-10>.
- Pfenninger, S. E., & Singleton, D. (2019). A critical review of research relating to the learning, use, and effects of additional and multiple languages in later life. *Language Teaching*, 52(4), 419-449. <https://doi.org/10.1017/S0261444819000235>.
- Sanjaya, A. (2018). The phonological analysis of a Japanese singer: The criticism of using songs in English teaching. *International Journal of Humanity Studies*, 2(1), 70-81. <https://doi.org/10.24071/ijhs.2018.020108>.
- Souza, H. K., Carlet, A., Jułkowska, I. A., & Rato, A. (2017). Vowel inventory size matters: Assessing cue-weighting in L2 vowel perception. *Ilha Do Desterro*, 70(3), 33-46. <https://doi.org/10.5007/2175-8026.2017v70n3p33>.
- Steele, J. (2009). Language experience in second language speech learning: In honor of James Emil Flege (review). *University of Toronto Quarterly*, 78(1), 151-153. <https://doi.org/10.1353/utq.0.0385>.
- van Leussen, J., & Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology*, 6, 1-12. <https://doi.org/10.3389/fpsyg.2015.01000>.
- Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50(1), 1-31. <https://doi.org/10.1017/S0261444816000276>.
- Yazawa, K., & Kondo, M. (2019). Acoustic characteristics of Japanese short and long vowels: Formant displacement effect revisited. *Proceedings of the 19th International Congress of Phonetic Sciences*. Melbourne, Australia.