# Formant Frequencies and Vowel Space Area in Javanese and Sundanese English Language Learners

ARUM PERWITASARI
*Leiden University Centre for Linguistics*
*Leiden Institute for Brain and Cognition*
*Leiden University*
*The Netherlands*
*a.perwitasari@hum.leidenuniv.nl*

MARIAN KLAMER
*Leiden University Centre for Linguistics*
*Leiden University*
*The Netherlands*

NIELS O. SCHILLER
*Leiden University Centre for Linguistics*
*Leiden Institute for Brain and Cognition*
*Leiden University*
*The Netherlands*

## ABSTRACT

*Several studies have documented how first language (L1) vowel systems play an important role in the vowel production of a second language (L2). L2 learners of Western languages who exhibit a smaller L1 vowel system are predicted to struggle with producing L2 vowels. However, there remains a paucity of evidence on how the L1 vowel system of non-Western languages interferes with L2 vowel production In this case, the focus is specifically on Javanese and Sundanese, two of the most widely spoken Indonesian local languages. This present study investigated how the six Javanese vowels and the seven Sundanese vowels influence the production of ten English vowels. In this experiment, 40 speakers, Javanese, Sundanese, and 10 native English speakers, participated in the production task. Spectral dimensions, including first (F1) and second formant (F2) frequencies, were analysed acoustically. According to the Speech Learning Model, Javanese and Sundanese speakers should have trouble producing similar vowels such as (/I, ɛ, ʊ/) and should not exhibit greater L2 differences with new vowels such as (/iː, æ, aː, ɔː, uː, ʌ, ɜː/). Indeed, the results demonstrated that the Javanese speakers did have different F1 and F2 values with the English vowels (/iː, æ, aː, ɔː, uː, ʌ, ɜː/) and the Sundanese speakers produced different F1 and F2 values for vowels (/æ, aː, ɔː/) when compared to the English native speakers. Interestingly, though vowels (/I, ʊ/) were considered to be similar vowels in the L1 vowel system, the Javanese and Sundanese speakers also showed differences in the formant structure. The vowel space area in the productions by Javanese and Sundanese speakers was slightly smaller than that of the native English speakers. The present study is expected to serve as a basis for future studies and provide the patterns of English vowels produced by Javanese and Sundanese learners of English.*

*Keywords: phonetics; language teaching; bilingualism; second language production; second language acquisition*

## INTRODUCTION

The effects of first language (L1) vowel systems in second language (L2) acquisition have been cross-linguistically assessed. In production tasks, if an L1 has a complex vowel system, the vowel space area in their speaking is predicted to be crowded (Flege 1995, 2003). Such crowded vowel space creates less room for new a vowel category and brings disadvantages in learning L2 vowels. This prediction, however, seems to be unresolved for an L1 vowel system with a small number of categories (Meunier et al. 2003). If an L1 has a small vowel system, the vowel space area may be less crowded as the L2 vowel sounds would easily

adjust into the same L1 category (Iverson & Evans 2007). The current study seeks to contribute to this assumption by examining the vowel space area between two regional languages in Indonesia, namely, Javanese and Sundanese. So far, the vowel space area in L2 vowel production by native speakers of Javanese and Sundanese has never been well investigated.

Vowel Space Area (VSA) measures spectral dimensions of formant frequencies, which are comprised of the vowel height or first formant (F1) and the degree of "backness", of the tongue, or second formant (F2), (Fant 1973, Gimson 1980, Cruttenden 2001). Formant frequencies are crucial to assess intelligibility and to identify the accuracy of pronunciation, as well as the naturalness of speech (Peterson & Barney 1952, Hillenbrand and Nearey 1999). L1 speakers who have a wide F1 range appeared to have higher intelligibility scores than the L1 speakers who have a narrow F1 range (Bradlow et al. 1996, Hazan & Markham 2004). The F2 range is, therefore, correlated with the intelligibility of words (Hazan and Markham 2004), not sentences (Bradlow et al. 1996).

The current study explores the VSA of Javanese and Sundanese English language learners by investigating the formant frequencies (F1 and F2 values) of English vowels. This paper first reviews and explains the literature and theories supporting such a prediction. It will then focus on the methods, results, and a discussion of the speech production experiment. We limit this study to the effects of the L1 vowel system; the consonantal context effects that may appear in this study are beyond the present scope of analysis.

## L2 SPEECH PRODUCTION MODEL

An often-discussed model rooted in the production of second language sounds is Flege's Speech Learning Model, SLM, (1995a, 1999, 2002). According to SLM, L2 learners can accurately produce L2 sounds if they have accurate understandings of L2 sound properties and the phonetic distance between L1 and L2 sounds. SLM hypothesizes that L2 learners would be less successful in learning *similar* sounds, indicated by a similar IPA symbol, in the L1 and L2 sounds with audible differences between the two (Flege 1997). The reason is that the similarity between the L1 and L2 sounds will block the formation of a new phonetic category. In contrast, L2 learners would experience no challenges in perceiving *new* sounds (Flege 1997). New L2 sounds, which are different from L1 categories and have no phonetic counterpart in the L1, enable the learners to develop new L2 categories (Flege 1997).

The phonetic and acoustic category differences of L1 Javanese, Sundanese and English motivated the present investigation. Javanese is a language which has an inventory of six vowel sounds with four allophonic pairs (/ɪ-i, ɛ-e, a, ə, ʊ-u, ɔ-o/) (Dudas 1976, Nothofer 2006, Clynes 1995, Wedhawati et al 2006), Sundanese has seven vowels (/ɪ, a, ə, ɨ, ɛ, ʊ, ɔ/) (Crothers 1978, Sudaryat, 2007) whereas English has a complex vowel system with ten monophthongs (/iː, ɪ, e, æ, ɑː, ɔː, ʊ, uː, ʌ, ɜː/) (Ladefoged 2001, 2006). The vowel diagram of Javanese, Sundanese and English is in Figure 1.

FIGURE 1. Articulatory plots for vowels of English (left), Javanese (middle) and Sundanese (right). The vowels in circles are considered as *similar* vowels between these three languages
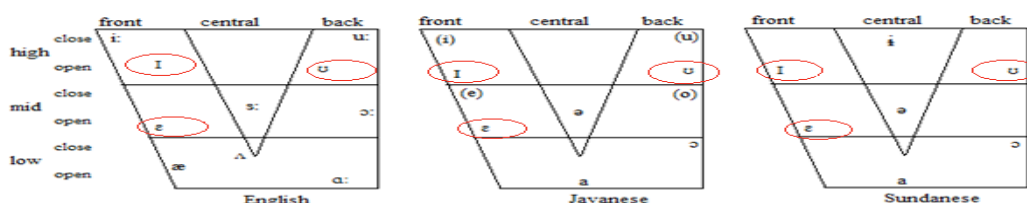
The inspection of a plot reveals that the position of English (/ɪ, ɛ, ʊ/) shows a similar IPA symbol and lies roughly in a similar vowel space area with the Javanese and Sundanese sounds. The English (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/), on the other hand, share a different plot and show different IPA symbols from the Javanese and Sundanese.

Cross-linguistic studies reveal the effects of an L1 vowel system with L2 vowel production. The L2 learners are predicted to use the cues from their L1 vowel system and apply them to L2 production. This situation may bring advantages for the L1 with a complex vowel system. Speakers with a larger L1 vowel system may be more successful in using assimilation by changing the L1 category representation to match the L2 vowels creating mergers or comprising categories (Flege 2003, MacKay et al. 2001). For instance, McAllister et al. (2002) found that English speakers with a larger vowel system performed better in producing Spanish vowels. Nevertheless, L2 vowel production is going to be problematic for L2 learners who have smaller L1 vowel systems than the target language. Iverson & Evans (2007) revealed that Germans and Norwegians, who have a complex L1 vowel system, were more accurate at recognising English vowels than French and Spanish speakers who have smaller L1 vowel systems.

A software package called *PRAAT* was used for the experiment of this study (Boersma & Weenink 2013). Generally, *PRAAT* allows for the recording of speech sounds and provides a visual representation of the acoustic signal through spectrograms, waveforms, and intonation contours (Olson 2014). The program can inspect the acoustic properties of utterances and show the formant frequencies of a sound as functions of time. *PRAAT* has allowed for the investigation of sounds, such as vowels and consonants, on speech segments. Almbark (2014) used *PRAAT* to examine the production of Standard Southern British English vowels by Syrian Arabic speakers and found that the learners differ from the native English speakers on duration, F1 and F2 values. Similarly, Shahidi et al (2012) analysed the production of velar stops of Malay and English by Malay speakers using PRAAT. Their results confirmed a clear contrast between voiced and voiceless velar stops of Malay and English.

To examine the production patterns of English vowels by Javanese and Sundanese English language learners, we set two research questions:

(1) Formant frequencies: would Javanese and Sundanese learners show the same formant frequencies producing new vowels (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/) in a native-like way? Would the production of similar vowels (/ɪ, ɛ, ʊ/) show different formant frequencies for the learners? Using SLM, it is predicted that the Javanese and Sundanese speakers may have difficulties for *similar* vowels (/ɪ, ɛ, ʊ/). However, Javanese and Sundanese speakers would be capable of producing *new* vowels (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/).

(2) Vowel space area: Would there be any influence of vowel space on production by their L1? If yes, how does the pattern impact the vowel space area? Based on the above findings that support the predictive role of the L1 vowel system and size of L2 vowel production (e.g. McAllister et al. 2002, Iverson and Evans 2007), Javanese and Sundanese speakers should be less successful at producing English vowels. With regard to the vowel space area, the Javanese and Sundanese learners should also produce a less crowded spectral dimension (Iverson & Evans 2007). Since previous research on Javanese and Sundanese learners' production of English is very scarce, we need to demonstrate the pattern of the vowel space area by the L2 learners compared to that of native English speakers.

## METHOD

### PARTICIPANTS

A total of 50 participants participated in the speech production experiment. Based on their first language background, they were divided into three groups: American English speakers (AmE), Javanese speakers (JE), and Sundanese speakers (SE).

The JE and SE participants all came from Central and West Java. They were students of various universities in Yogyakarta. The mean age for the 20 Javanese speakers (JE groups) was 21.9 (Age range: 20 - 30 years, SD = 1.16) and for the 20 Sundanese speakers (SE groups) it was 22 (Age range: 21 – 32 years, SD = 2.41). The average age at which the JE and SE speakers began learning English was 8.7 years old (JEL: SD = 1.5; SEL: SD = 2.5). The JE and SE speakers had no experience in travelling abroad. They were proficient in their first language in the sense that they were still using their L1 in their daily lives. The JE and SE participants had learnt English for a minimum of 9 years in formal education (JE: M= 11.75 years, SD = 2.2; SE: M = 9.8 years, SD = 2.4).

The 10 English speakers (AmE group) had come from the central and western areas of the United States. Their mean age was 26.2 (Age range: 23 – 36 years, SD = 1.75). All participants were tested at either the Gadjah Mada University in Yogyakarta or the Padjajaran University in West Java, both in Indonesia. Participants were provided with a consent form.

### STIMULI

In this study, only English monophthongs were taken into consideration. The Javanese (Jav), Sundanese (Sun), and American English (AmE) speakers produced 10 English monophthongs (/iː, ɪ, e, æ, ɜː, ʌ, ɑː, ɔː, ʊ, uː/). The vowels were embedded in two different consonantal contexts, namely in /bVd/ and /hVd/ syllables. The ten monophthongs in the /bVd/ context were *bead, bid, bed, bad, bird, bud, body, bawd, Buddhist, and booed*. In the /hVd/ context, the words included *heed, hid, head, had, heard, hudd, hod, hawed, hood, and who'd* (Ladefoged 2001). All syllables were embedded in the sentence frame "*I say (bVd)/ (hVd) again*". Sentences were presented in letters on a computer screen. During the recording, subjects repeated the sentences twice.

### PROCEDURE

Before the production experiment took place, all participants completed a brief sociolinguistics questionnaire and a consent form. The first part of the questionnaire elicited their demographic information and their experiences with both their native and second languages. The subjects reported their parents' first language and how often they used their native language. They stated their choice of language at home and at school. The second part of the questionnaire elicited the subjects' background in the second language. The non-native speakers subjects reported their ages when they began learning English and how long they had been studying English. They listed the second language that they have learnt and the competence level of the second language. The Javanese and Sundanese subjects mostly listed Arabic and Japanese as other non-native languages that they have learnt. They shared how long they had learnt English and they confirmed that they had never lived in any English speaking countries.

All participants received a short introduction monologue which contained words simulated for the recording. The introduction monologue was presented on a computer screen. In the meantime, the researcher explained about the experiment and the recording

procedures that would be involved. The subjects were given as much time as they needed and were encouraged to ask and comment at any point during the explanation. To start, each of the subjects sat in front of a computer display with active mode recording tools (audio and video recorders, as well as a headset microphone). Once the stimuli appeared on the screen, subjects produced the sentence, for instance "*I say bad again*". All of the stimuli were presented twice: random and sequenced orders. Their speech production was documented and stored in a computer file. Both audio and video recordings were handled carefully and used for acoustic analysis.

All participants were recorded in a sound-attenuated room. Items were digitised using a digital audio recorder (H4N Zoom) and an adjustable microphone headset (Sennheiser PC 141) with a 44.1 kHz/16 bit sampling. The distance that the microphone was set away from the speaker's mouth was approximately 3 cm to create a constant sound record for the entire session for every subject. After the completion of each experiment, subjects were given a post-experiment questionnaire. The post-experiment questionnaire was given to obtain information with regard to the subjects' experiences in producing the stimuli that they received in the experiments. Afterwards, they received their compensation and were allowed to share any concerns about the experiment in a written form.

## ANALYSIS

The study utilised Praat 5.3.56 (Boersma & Weenink, 2013) for annotating speech. The Javanese group produced 800 English vowels (2 contexts x 20 speakers x 10 vowels x 2 repetitions), the Sundanese group produced 800 English vowels (2 contexts x 20 speakers x 10 vowels x 2 repetitions) and the American English group produced 400 English vowels (2 contexts x 10 speakers x 10 vowels x 2 repetitions). The total corpus of data amounts to 2000 vowels. Formant frequencies of the participants' speech were tracked through the estimation and plot of each vocal tract. For pitch and formant frequencies (F1 and F2), the values were traced by identifying the formant peak at a chosen time point. The values of the F1 and F2 were specifically measured at the midpoint of the steadily stated selected vowel. The F1 and F2 values were then converted to the Bark scale using the following formula: $Zi = 26.81/(1+1960/Fi)- 0.53$ (Traunmüller 1988).

For statistical analysis, we performed a series of repeated ANOVA test measures using SPSS 22.0 (IBM 2013). A repeated measures analysis was conducted to examine the effect of the factored VOWEL and CONTEXT on the first formants (F1) and second formant (F2) frequencies. If Mauchly's test is significant, we used the Greenhouse Geisser prediction for the entire analysis. The independent variables in this study were L1 GROUP (Javanese, Sundanese, English), VOWEL (/iː, ɪ, e, æ, ɜː, ʌ, ɑː, ɔː, ʊ, uː/) and CONTEXT (/bVd/ vs /hVd/). The dependent variables were the F1 and F2 values. To follow the differences between Javanese vs. English and Sundanese vs. English, we conducted a series of independent t-test analysis to determine whether there was a statistically significant difference between the means of F1 and F2 in two groups (JE vs. AmE and SE vs. AmE).

## RESULTS

### FORMANT FREQUENCIES OF JAVANESE VS ENGLISH SPEAKERS

A repeated ANOVA measure with a Greenhouse-Geisser correction determined that the first formant frequency (F1) was significantly affected by the VOWEL [$F(4.4, 123) = 61.39$, $p < .001$] and CONTEXT, [$F(1, 28) = 53.8$, $p < .001$]. Significant effects were noted for the

factors in the VOWEL x GROUP, [$F$(4.4, 123) = 2.5, p < .05] and the factors in the CONTEXT x GROUP [$F$(1, 28) = 4.9, p < .05]. There was a significant difference between the factors in VOWEL X CONTEXT, [$F$(4.8, 135.4) = 71.1, p < .001] and the speaker groups on the factors  VOWEL x CONTEXT x GROUP [$F$(4.8, 135.4) = 5.9, p < .001].

The second formant frequency (F2) was significantly affected by both the VOWEL [$F$(4.3, 121) = 40.6, p < .001] and the CONTEXT [$F$(1, 28) = 120.8, p < .001]. There was an interaction of the factors in CONTEXT x GROUP, [$F$(1, 28) = 5.02, p < .05]. However, no effect was seen for the factors in VOWEL x GROUP [$F$(4.3, 121) = 1.8, p = .128]. Significant effects were noted on the factors in VOWEL x CONTEXT, [$F$(3.9, 110.6) = 78.9, p < .001], but no effect was seen for the factors of the VOWEL x CONTEXT x GROUP [$F$(3.9, 110.6) = 2.3, p = .063].

To test whether the F1 and F2 means were significantly different from each other, an independent t-test was conducted. The t-test results showed the following pattern (see Table 1).

TABLE 1. Independent t-test results comparing Javanese (JAV) and American English (Eng) speakers on first formant (F1) and second formant (F2) frequencies of the ten English vowels produced in the /bVd/ and /hVd/ contexts, * = p < .05, ** = p < .001

| English Vowel | Context | F1 (Hz) | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | L1 Eng | L1 Jav | t | df | P |
| /iː/ | /bVd/ | 314 | 416 | -2.087 | 28 | 0.046* |
| /ɜː/ | | 492 | 514 | -0.619 | 28 | 0.541 |
| /ɑː/ | | 795 | 639 | 2.688 | 28 | 0.012* |
| /ɔː/ | | 724 | 657 | 1.289 | 28 | 0.046* |
| /ʌː/ | | 374 | 389 | -0.052 | 28 | 0.607 |
| /ɪ/ | | 482 | 403 | 2.338 | 28 | 0.027* |
| /e/ | | 610 | 578 | 0.685 | 28 | 0.499 |
| /æ/ | | 752 | 635 | 2.132 | 28 | 0.042* |
| /ʌ/ | | 639 | 642 | -0.04 | 28 | 0.968 |
| /ʊ/ | | 376 | 453 | -1.494 | 28 | 0.027* |
| /iː/ | /hVd/ | 345 | 388 | -1.515 | 28 | 0.141 |
| /ɜː/ | | 509 | 540 | -0.758 | 28 | 0.454 |
| /ɑː/ | | 788 | 657 | 2.129 | 28 | 0.042* |
| /ɔː/ | | 750 | 666 | 1.492 | 28 | 0.147 |
| /ʌː/ | | 358 | 400 | -1.247 | 28 | 0.223 |
| /ɪ/ | | 489 | 433 | 1.738 | 28 | 0.093 |
| /e/ | | 616 | 603 | 0.214 | 28 | 0.832 |
| /æ/ | | 797 | 659 | 2.253 | 28 | 0.032* |
| /ʌ/ | | 641 | 599 | 0.618 | 28 | 0.032* |
| /ʊ/ | | 512 | 455 | 0.818 | 28 | 0.420 |
| English Vowel | Context | F2 (Hz) | | | | |
| | | L1 Eng | L1 Jav | t | df | P |
| /iː/ | /bVd/ | 2407 | 2342 | 0.517 | 28 | 0.609 |
| /ɜː/ | | 1607 | 1541 | 1.02 | 28 | 0.316 |
| /ɑː/ | | 1393 | 1542 | -2.055 | 28 | 0.049* |
| /ɔː/ | | 1275 | 1288 | -0.184 | 28 | 0.856 |
| /ʌː/ | | 1281 | 1211 | 0.831 | 28 | 0.413 |
| /ɪ/ | | 1837 | 2133 | -2.222 | 28 | 0.035* |
| /e/ | | 1849 | 1942 | -0.941 | 28 | 0.355 |
| /æ/ | | 1792 | 1868 | -0.813 | 28 | 0.423 |
| /ʌ/ | | 1584 | 1582 | 0.021 | 28 | 0.984 |
| /ʊ/ | | 1331 | 1425 | -1.215 | 28 | 0.235 |
| /iː/ | /hVd/ | 2387 | 2329 | 0.473 | 28 | 0.640 |
| /ɜː/ | | 1639 | 1618 | 0.265 | 28 | 0.793 |
| /ɑː/ | | 1413 | 1444 | -0.317 | 28 | 0.754 |
| /ɔː/ | | 1362 | 1358 | 0.058 | 28 | 0.954 |
| /ʌː/ | | 1245 | 1254 | -0.117 | 28 | 0.908 |
| /ɪ/ | | 1880 | 2086 | -1.775 | 28 | 0.090 |
| /e/ | | 1856 | 1945 | -0.768 | 28 | 0.439 |
| /æ/ | | 1807 | 1890 | -1.051 | 28 | 0.302 |
| /ʌ/ | | 1624 | 1460 | 2.065 | 28 | 0.048* |
| /ʊ/ | | 1533 | 1290 | 2.54 | 28 | 0.017* |

Independent t-tests showed that there was a significant difference between the Javanese and American speakers on F1 values for English vowels (/iː, ɑː, ɔː, ɪ, ʊ, æ/) in the /bVd/ syllable and vowels (/ɑː, æ, ʌ/) in the /hVd/ syllable. The Javanese and American English speakers showed a significant difference in the F2 values for English vowels (/ɑː, ɪ/) in the /bVd/ syllable and vowels (/ʌ, ʊ/) in the /hVd/ syllable.


FORMANT FREQUENCIES OF SUNDANESE VS ENGLISH SPEAKERS


A repeated ANOVA measure analysis with a Greenhouse-Geisser correction test was conducted to compare the main effects of the VOWEL and CONTEXT factors on the F1 and F2 values between the Sundanese and American English groups. The F1 values differed significantly for the VOWEL factor [$F$(3.2, 90.8) = 41.1, p < 0.001] and CONTEXT [$F$(1,

28) = 70.7, p < 0.001]. The results indicated that there was an interaction between the VOWEL x GROUP factors [$F$(3.2, 90.8) = 3.04, p < 0.05] and the CONTEXT x GROUP factors [$F$(1, 28) = 6.4, p < 0.05]. The interaction was significant for the VOWEL x CONTEXT factors [$F$(4.6, 130) = 695.1, p < 0.001] and the VOWEL x CONTEXT x GROUP factors [$F$(4.6, 130) = 6.0, p < 0.001].

The F2 values differed significantly for the VOWEL factor [$F$(4.2, 117) = 34.6, p < 0.001] and CONTEXT [$F$(1, 28) = 84.3, p < 0.001]. Significant effect was seen for the VOWEL x GROUP factors [$F$(4.2, 117) = 34.6, p < 0.05] and CONTEXT x GROUP, [$F$(1, 28) = 6.2, p < 0.05]. Significant effects were noted for the VOWEL x CONTEXT factors [$F$(4, 113.7) = 71.5, p < 0.001], and the VOWEL x CONTEXT x GROUP factors [$F$(4, 113.7) = 3.3, p < 0.05]. An independent t-test was conducted to examine the difference between the Sundanese and American English speakers on F1 and F2 values (see Table 2).

TABLE 2. Independent t-test results comparing Sundanese (SUN) and American English (Eng) speakers on first formant (F1) and second formant (F2) frequencies of the ten English vowels produced in the /bVd/ and /hVd/ contexts, * = p < .05, ** = p < .001

| English Vowel | Context | F1 (Hz) | | | | |
|---|---|---|---|---|---|---|
| | | L1 Eng | L1 Sun | T | df | P |
| /iː/ | /bV d/ | 314 | 388 | -1.978 | 28 | 0.058 |
| /ɜː/ | | 492 | 480 | 0.534 | 28 | 0.598 |
| /ɑː/ | | 795 | 673 | 3.824 | 28 | 0.001** |
| /ɔː/ | | 724 | 650 | 1.287 | 28 | 0.209 |
| /uː/ | | 374 | 380 | -0.188 | 28 | 0.209 |
| /ɪ/ | | 482 | 370 | 4.349 | 28 | 0.000** |
| /e/ | | 610 | 560 | 1.287 | 28 | 0.209 |
| /æ/ | | 752 | 618 | 2.889 | 28 | 0.007* |
| /ʌ/ | | 639 | 628 | 0.231 | 28 | 0.819 |
| /ʊ/ | | 376 | 463 | -1.916 | 28 | 0.066 |
| /iː/ | /hV d/ | 345 | 360 | -0.605 | 28 | 0.550 |
| /ɜː/ | | 509 | 526 | -0.421 | 28 | 0.677 |
| /ɑː/ | | 788 | 680 | 2.489 | 28 | 0.019* |
| /ɔː/ | | 750 | 634 | 2.136 | 28 | 0.042* |
| /uː/ | | 358 | 405 | -1.05 | 28 | 0.303 |
| /ɪ/ | | 489 | 408 | 2.634 | 28 | 0.014* |
| /e/ | | 616 | 618 | -0.052 | 28 | 0.959 |
| /æ/ | | 797 | 669 | 2.146 | 27 | 0.041* |
| /ʌ/ | | 641 | 688 | -0.864 | 28 | 0.395 |
| /ʊ/ | | 512 | 402 | 3.72 | 28 | 0.001* |
| **English Vowel** | **Context** | **F2 (Hz)** | | | | |
| | | L1 Eng | L1 Sun | T | df | P |
| /iː/ | /bV d/ | 2407 | 2399 | 0.056 | 28 | 0.955 |
| /ɜː/ | | 1607 | 1517 | 1.294 | 28 | 0.206 |
| /ɑː/ | | 1393 | 1589 | -2.263 | 28 | 0.032* |
| /ɔː/ | | 1275 | 1306 | -0.312 | 28 | 0.757 |
| /uː/ | | 1281 | 1235 | 0.377 | 28 | 0.709 |
| /ɪ/ | | 1837 | 2365 | -4.569 | 28 | 0.000** |
| /e/ | | 1849 | 1864 | -1.959 | 28 | 0.060 |
| /æ/ | | 1792 | 1961 | -2.202 | 28 | 0.036* |
| /ʌ/ | | 1584 | 1626 | -0.381 | 28 | 0.706 |
| /ʊ/ | | 1331 | 1396 | -0.815 | 28 | 0.422 |
| /iː/ | /hV d/ | 2387 | 2294 | -0.192 | 28 | 0.490* |
| /ɜː/ | | 1639 | 1610 | 0.343 | 28 | 0.734 |
| /ɑː/ | | 1413 | 1459 | -0.515 | 28 | 0.611 |
| /ɔː/ | | 1362 | 1441 | 0.982 | 28 | 0.335 |
| /uː/ | | 1245 | 1205 | 0.362 | 28 | 0.720 |
| /ɪ/ | | 1880 | 2275 | -2.625 | 28 | 0.014* |
| /e/ | | 1856 | 1976 | -1.031 | 28 | 0.311 |
| /æ/ | | 1807 | 1828 | -0.162 | 28 | 0.872 |
| /ʌ/ | | 1624 | 1614 | 0.099 | 28 | 0.921 |
| /ʊ/ | | 1533 | 1280 | 2.908 | 28 | 0.007 |

Independent t-tests showed that there was a significant difference between the Sundanese and American speakers in the values of the F1 for English vowels (/ɑː, ɪ, æ/) in the /bVd/ syllable and vowels (/ɑː, ɔː, ɪ, æ, ʊ/) in the /hVd/ syllable. The English vowels (/ɑː, ɪ, æ/) in the /bVd/ context and (/iː, ɪ/) in the /hVd/ context differed significantly between the Javanese and American English speakers.

## VOWEL SPACE AREA

To compare the vowel space area of the Javanese, Sundanese, and American English speakers in the /bVd/ context, the F1 and F2 values were inserted into a vowel quadrangle table (see Figure 2). The vowel quadrangle table refers to the place of articulation in the mouth and represents the position of the tongue for each vowel.

FIGURE 2. Distribution of the F1 and F2 (in Bark) in the English vowels as produced by L1 English, L1 Javanese, and L1 Sundanese speakers in the /bVd/ context
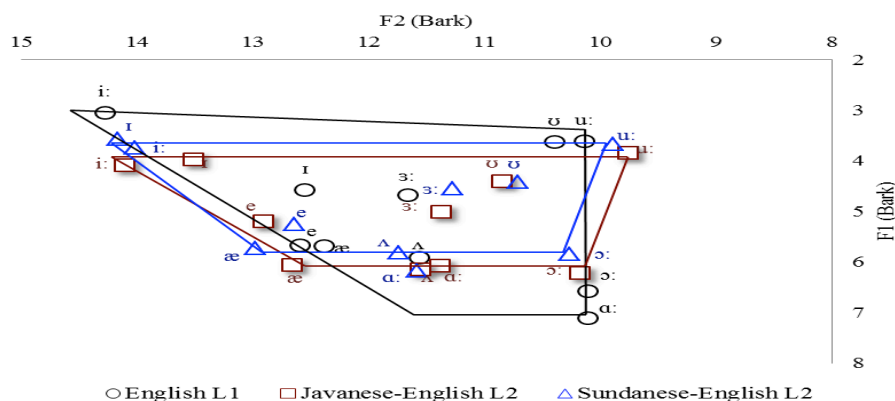


Figure 2 is a visualisation of the vowel space area of the English vowels produced by the Javanese and Sundanese speakers, as well as the English speakers. From the overall formant values, some investigation could be taken:

1.  The Javanese and Sundanese speakers had a slightly smaller vowel space than the native English speakers.
2.  The vowel /i:/ produced by the Javanese and Sundanese L2 learners was spectrally lower and further back than the vowel /i:/ by the native English speakers. In contrast, the Javanese and Sundanese L2 learners produced a higher and more anterior vowel /I/ than the L1 English speakers.
3.  The production of the vowel /æ/ by the Javanese native speakers was lower than that of the Sundanese native speakers and native English speakers.
4.  There is no overlap in the vowel /ɜ:/ in which the Javanese and Sundanese L2 learners produce the vowel lower and further back than the L1 English speakers.
5.  The L1 Javanese and Sundanese speakers produced the vowel /ɑ:/ higher and in a more anterior location than the native English speakers. However, their production of the vowel /ɑ:/ fell close to the production of the vowel /ʌ/.
6.  The Javanese and Sundanese speakers production for the vowel /ɔ:/ showed a slight disparity. They produced a slightly higher vowel /ɔ:/ than the native English speakers.

FIGURE 3. Distribution of the F1 and F2 (in Bark) in English vowels as produced by L1 English, L1 Javanese, and L1 Sundanese speakers in the /hVd/ context



148
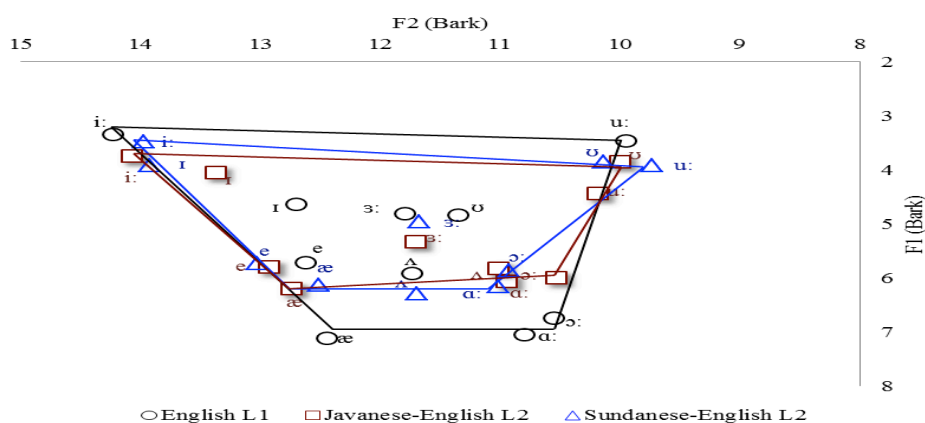
To clearly examine the effect of the consonantal context, the vowel spaces of the F1 and F2 values produced by the Javanese and Sundanese L2 learners in the /hVd/ context was conducted (see Figure 3). The figure allows for a number of investigations:

1.  The Javanese and Sundanese speakers produced a vowel space, which was slightly smaller than that of the native English speakers.
2.  The quadrangle table showed that the vowel /iː/ of the learners was slightly further back than that of the native English speakers. However, the learners' production of the vowel /ɪ/ was higher and more anterior than the vowel /ɪ/ produced by the L1 English speakers. The Sundanese speakers produced the vowel /ɪ/ relatively closer to their production of the vowel /iː/.
3.  The Javanese and Sundanese speakers articulated the vowel /ʊ/ more to the back and closer to the production of the vowel /uː/ than the vowel /ʊ/ produced by the L1 English speakers. The Javanese speakers produced it further back than both the Sundanese speakers and the native English speakers.
4.  The Javanese speakers produced the vowel /ʌ/ spectrally adjacent to their L2 production of the vowels /ɔː/ and /ɑː/.
5.  Both groups of L2 learners produced the vowel /æ/ with a more open articulatory height and more closely similar to the vowel /e/ by the L1 English speakers.

## DISCUSSION

Two main research questions motivated the current study. Firstly, can Javanese and Sundanese learners produce new vowels (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/) in a native-like way? Will similar vowels (/ɪ, ɛ, ʊ/) create difficulties for the learners? If the SLM (Flege 1995a, 1999, 2002) successfully predicts L2 vowel production, Javanese and Sundanese speakers may not exhibit differences in formant structure for *new* vowels (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/), but may show different formant frequencies for *similar* vowels (/ɪ, ɛ, ʊ/). Contrary to our prediction, the *new* vowels (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/) to the L2 learners were considerably more difficult to produce. The Javanese speakers showed different formant frequencies for the *new* vowels (/iː, æ, ɑː, ɔː, ɜː/) but *similar* vowels (/ɪ, ʊ/) in the /bVd/ context. They also demonstrated significantly different formant frequencies for the *new* vowels (/ʌ, ɑː, æ/) and *similar* vowel (/ʊ/) in the /hVd/ context. Unlike the Javanese speakers, Sundanese speakers had significantly different formant frequencies for only *similar* vowel (/ɪ/) and *new* vowels (/æ, ɑː/) in the /bVd/ context, and similar vowel (/ɪ/) and new vowels (/ɑː, ɔː, iː/) in the /hVd/ context.

The second research question relates to the vowel space area of the Javanese and Sundanese learners of English and whether or not there will be any difference in the *pattern* of the vowel space area. We predicted that because the L1 Javanese and Sundanese have a smaller vowel system, the learners might find difficulties and show a less crowded vowel space area in producing English vowels. The visual investigation of the vowel space area suggested that L2 vowel production is problematic for L2 learners who have a smaller L1 vowel system than that of the target language. The vowel space area in L2 production confirms that the L2 learners struggled to attune with the formant frequencies of L2 sounds. The Javanese and Sundanese speakers, who have smaller L1 vowel systems than that of English speakers, did not make accurate formant structures when compared to the native English speakers. The findings support the predictive role of the L1 vowel system and size in L2 vowel production as previously shown in Iverson & Evans (2007). The Javanese and

Sundanese speakers, as predicted, are less successful at producing English vowels than the native English speakers.

Considering the evidence that the size of the L1 vowel system may predict the accuracy of L2 production, (Iverson and Evans 2007), it may be taken as evidence that the Javanese and Sundanese learners of English in this study did not have a fully developed category for English vowels. The difficulties experienced by the Javanese and Sundanese speakers are mostly shown in the F1 values indicating that they had different vowel height realisations, rather than in the degree of *backness*, when compared to the English speakers. However, as already mentioned, the speakers recruited in this study were university students studying English. Although the proficiency level of the Javanese and Sundanese speakers in the current study is comparable, their production accuracy for the English vowels showed differences. Presently, we can only speculate that the causes may be due to the first language vowel system. Javanese has (/ɪ-i, ɛ-e, ʊ-u, ɔ-o/) allophonic pairs (Dudas 1976, Nothofer 2006, Clynes 1995, Wedhawati et al 2006). Besides, Javanese and Sundanese L1 were reported to have no vowel length contrast (van Zanten and van Heuven 1997) and exhibit no long vowels and diphthongs (Gordon 2006). Another factor may also be predicted by explicit phonetic training (see Evans & Iverson 2009, Wang & Munro 1999).


## CONCLUSION


To the best of our knowledge, this is the first study to examine both the formant frequencies and the vowel space area of Javanese and Sundanese learners of English. We found that the production of English vowels was challenging for the Javanese and Sundanese EFL learners not only for *similar* sounds such as (/ɪ, ɛ, ʊ/), but also for *new* sounds such as (/iː, æ, ɑː, ɔː, uː, ʌ, ɜː/). The Javanese and Sundanese native speakers created random patterns based on the degree of difficulty. Moreover, the Javanese and Sundanese L2 learners differed, to some extent, in the vowel space area of the English vowels compared to the native speakers. The quadrangle table containing F1 and F2 values revealed detailed information about the deviations. The Javanese and Sundanese speakers showed a smaller vowel space area than that of the native English speakers.

The current study will serve as a basis for future studies and provide the vowels produced by Javanese and Sundanese learners of English. With respect to pedagogical implications, the results enabled the L2 learners of English to identify pronunciation accuracy and helped to increase intelligibility (Peterson & Barney 1952, Hillenbrand & Nearey 1999). As highlighted by Gass & Selinker (2008), the gaps in vowel space area would cause the L2 learners to notice the divergent areas of L2 sounds. Noticing the gaps between the L2 learners and the native speakers' production would increase the learner's awareness and increase the L2 competency (Munro & Derwing 2008, Gass & Selinker 2008).

In future research, given the differences in formant frequencies, it would be worth exploring the similar pattern that may or may not be apparent in the L2 learners production of different languages and consonantal contexts. More data on cross-linguistic comparisons would make it possible to examine the differences and similarities of vowel space area. Acoustic investigations comparing the vowel space area of other Indonesian regional languages would provide a heightened awareness of L2 learners with various dialects or language backgrounds.

REFERENCES

Almbark, R. (2014). Perception and production of vowels by foreign language learners: Towards a foreign language model. *Proceedings of the International Symposium on the Acquisition of Second Language Speech*. Concordia Working Papers in Applied Linguistics, 5.

Bradlow, A., Torretta, G., and Pisoni, D. (1996). Intelligibility of normal speech 1: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication. Vol. 20,* 255 – 272.

Boersma, P., & Weenink, D. (2013). *Praat: doing phonetics by computer [Computer program]* Version 5.3.51, retrieved 5 July 2013 from htt[://www.praat.org/.

Clynes, A. (1995). Javanese. In Tryon D.T. (Ed.) *An Introduction to Austronesian Studies 1: Fascicle 1: Comparative Austronesian Dictionary*. Berlin, New York: Mouton de Gruyter, 469-484.

Crothers, J. (1978). Typology and universals of vowel systems. In J.H. Greenberg (Ed.), *Universals of human language,* Vol 2 Phonology (pp. 94-152). Stanford, California, Stanford University Press.

Cruttenden, A. (2001). *Gimson's Pronunciation of English*. 5th ed. London: Edward Arnold.

Dudas, K. (1976). *The Phonology and Morphology of Modern Javanese*. Ann Arbor, Mich.: University Microfilms.

Fant, G. (1973). *Speech Sounds and Features.* MIT Press: Cambridge.

Flege, J. E. (1995). Two methods for training a novel second-language phonetic contrast. *Applied Psycholinguistics. Vol. 16,* 425-442.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 229-273). Timonium, MD: York Press.

Flege, J. E. (1995a). The phonetic study of bilingualism. In G. Bloothooft, V. Hazan & J. Llisterri (Eds.), *European Studies in Phonetics and Speech Communication* (pp. 98-103). Utrecht: OTS publication.

Flege, J. E. (1997). English vowel production by Dutch talkers: More evidence for the "new" vs "similar" distinction. In A. James & J. Leather (Eds.), *Second-language Speech. Structure and Process* (pp. 11-52). Berlin: Mouton de Gruyter.

Flege, J. E. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis* (pp. 101-132). Hillsdale, NJ: Lawrence Erlbaum.

Flege, J.E., MacKay, I. & Meador, D. (1999). Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America. Vol. 106,* 2973-2987.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske and A. Rohde (Eds.), *An integrated view of language development papers in honor of henning wode* (pp. 217- 244). Trier: Wissenschaftlicher Verlag.

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

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

Gass, M. & Selinker, L. (2008). *Second Language Acquisition: An Introductory Course* (3rd ed.). London: Sage.

Gimson, A. C. (1980). *An Introduction to the Pronunciation of English (3rd Ed.),* pp xvi + 352. London: Edward Arnold.

Gordon, M. (2006). *Syllable Weight: Phonetics, Phonology, Typology*. London: Routledge.

Hazan, V., & Markham, D. (2004). Acoustic-phonetic correlates of talker intelligibility for adults and children. *Journal of the American Academy of Audiology. Vol. 116(5),* 3108–3118.

Hillenbrand, J. M., & Nearey, T. M., (1999). Identification of resynthesized /hvd/ utterances: Effects of formant contour. *Journal of the Acoustical Society of America. Vol. 406 (6),* 3509–3523.

IBM Corp. (2013). *IBM SPSS Statistics for Macintosh*, Version 22.0. Armonk, NY: IBM Corp.

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

Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Languages*. Blackwell.

Ladefoged, P. (2006). *A Course in Phonetics*. 5th Ed. Boston: Thomson Wadsworth.

MacKay, I. R. A., Meador, D., & Flege, J. E. (2001). 'The identification of English consonants by native speakers of Italian. *Phonetica. Vol.* 58, 103– 125.

McAllister, R., Flege, J., & Piske, T. (2002). The Influence of L1 on the Acquisition of Swedish quantity by Native Speakers of Spanish, English, Estonia. *Journal of Phonetics. Vol.* 30, 229-258.

Meunier, C., Frenck-Mestre, C., Lelekov-Boissard, T. & Le Besnerais, M. (2003). Production and perception of foreign vowels: Does the density of the system play a role?" in the *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain.

Munro, M. J. & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning. Vol.* 58, 479-502.

Nothofer, B. (2006). Javanese. In. Brown, K & Ogilvie, S. (Eds.). *Concise Encyclopedia of Languages of the world* (pp. 560-561). Oxford: Elsevier Ltd.

Olson, D.J. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning and Technology. Vol.* 189(3), 173-192. Retrieved from http://llt.msu.edu/issues/octiber2014/olson.pdf

Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America. Vol.* 24, 175–184.

Shahidi, A.H, Rahim Aman, & Ab. Samad Kechot (2012). Production and Perception of English Word Final Stops by Malay Speakers. *GEMA Online® Journal of Language Studies. Vol.* 12(4), 1109 - 1125.

Sudaryat, Y. (2007). *Elmuning Sora Basa Sunda*. Bandung: Pustaka Luang Bandung.

Traunmüller, H. (1988). Paralinguistic variation and invariance in the characteristic frequencies of vowels. *Phonetica. Vol.* 45, 1-29. Doi: 10.1159/000261809.

Van Zanten, E., & Van Heuven, V.J. (1997). Effects of word length and Substrate language on the temporal Organisation of words in Indonesian. In: C. Ode and W.A.L. Stokhof (Eds.), *Proceedings of the 7th International Conference on Austronesian Linguistics* (pp. 201-16). Amsterdam/Atlanta: Rodopi.

Wang, X., & Munro, M.J. (1999). The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention on temporal and spectral cues. *Proceedings of the 14th International Congress of Phonetic Sciences*: 125-128.

Wedhawati, Wiwin Erni Siti Nurlina, Edi Setiyanto, Marsono, Restu Sukesi, & I. Praptomo Baryadi. (2006). *Tata Bahasa Jawa Mutakhir*. Yogyakarta: Penerbit Kanisius.