



Rudha Widagsa

got his Master of Arts in Linguistics from Gadjah Mada University, Indonesia. He currently works at Faculty of Teacher Training and Education, University of PGRI Yogyakarta (UPY), Indonesia. His research interests are second language acquisition, second language phonetics, applied linguistics and teaching English as a foreign language.

1-10

Closing Diphthongs among Javanese Learners of English

ABSTRACT

This paper describes adult acquisition of L2 English diphthongs among Javanese learners of English. The objectives of this study are concerned with whether vowels are more or less diphthongal by measuring the change of first formant (F1) and speech duration to examine the influence of Javanese (L1) phonetic system on the acquisition of English as a second language. The acoustical analysis was conducted by comparing the formant frequencies and duration of L2 with the corresponding diphthongs by native speakers of English (NE). The purpose was to test the hypothesis that difficulty in acquiring an L2 contrastive category is related to the role in the L1 of the phonetic feature upon which the L2 category is based. Although F1 values of L2 differed in the start point of closing diphthongs, there was only one closing diphthong /QŠ/ produced by L2 which was statistically different from L1 based on the F1 changes. Surprisingly, the diphthong /Tj/ showed an increase F1 value instead of decrease which means that the Javanese production was more open. The result in duration showed that there were crucial differences in diphthongal duration produced by Javanese learners of English. Javanese tended to lengthen the diphthongal words.

Keywords: diphthong, experimental phonetics, second language acquisition

INTRODUCTION

In English as second language (ESL) teaching, learners of English particularly adult learners, frequently make mistakes even failure in producing native like pronunciation because their L1 highly influences L2 production. Admittedly, as a speaker of a language gets older, mother tongue interference cannot simply be

dislodged. Adults tend to stabilize their language learning at a certain stage. Ellis (1994, p.09) stated that learner speech was 'structurally organized' in the sense that it constituted a system in its own right. This process called fossilization; when a sound in L2 consistently replaced with a sound which is phonetically close to L1. Thus, the speech production in L2 is different from the similar speech produced by a native speaker. Selinker (1972) also noted that relatively few adult learners reach native-speaker competence. Contrastive analysis hypothesis (CAH) proposed by Lado (1957, p.2) was also based on the assumption that people who comes into contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to their native language will be simple for him, and those elements that are different will be difficult.

Likewise, Javanese learners of English frequently have difficulties with uncommon English vowel system because there are dissimilarities in articulating vowel sounds of these two sound systems. In L2 learning, the level of difficulty experienced by the learner will be directly related to the degree of linguistic difference between L1 and L2. Difficulties will manifest themselves in errors; the greater the difficulty, the more frequent the errors (Ellis, 1994, p.308). Similarly, the difference between English and Javanese vowel system could be a barrier for Javanese learners of English who want to sound more native-like.

Javanese and English have vast differences in vowel system particularly diphthongs. Javanese phonology includes 8 vowel phonemes: 6 phonemes and 2 allophonic pairs (Uhlenbeck, 1963). An Indonesian linguist, Marsono (1999, pp.55-58), reported 10 vowels (including allophonic variants) and

5 diphthongs in Javanese. English, on the other hand, has a large number of diphthongal sounds which consist of a movement or glide from one vowel to another. One of the most common pronunciation mistakes that result in a learner of English having a "foreign" accent is the production of pure vowels where a diphthong should be pronounced (Roach, 1998, p.18).

Each vowel including diphthong has a unique structure of formant frequency which indicates the quality of the vowel. Formants have long been known as suited parameters for describing vowel production due to its correlation with traditional articulatory transcription of vowel (Hawkins & Midgley, 2005). The relationship between F1 and F2 determines the acoustic quality of the vocoid articulations. The longer the cavity, the lower F1 will be; the shorter it is, the higher F1 will be. This correlates with tongue-height; the pharyngeal tube continues into the mouth and is narrowed in varying degrees by the movement up and down of the tongue. Similarly, F2 correlates with the length of the oral cavity in terms of frontness and backness of the tongue body. The longer the front cavity, the lower F2 will be (= back vowels); the shorter it is, the higher F2 will be (= front vowels) (Lodge, 2009, pp. 190-199).

Thus, formant frequencies can be used to measure the accuracy of vowel production. Nevertheless, in Indonesia, English pronunciation instruction mainly focuses on audio-lingual approach such as minimal pair drills and articulatory descriptions. Correction and analysis of students' production generally depends on the teachers' listening judgment. In fact, one of the accurate ways to discern the differences would be based on acoustic analysis. Therefore, this research attempts to investigate

L2 production in English pronunciation especially diphthongs by examining the acoustic difference between L2 diphthongs and native speakers (NE) diphthongs.

This study investigates pronunciation of L2 diphthongs and attempts to answer two questions (1) to what extent is the pronunciation of L2 diphthongs different from the corresponding diphthong in NE? (2) is there any difference in speech duration between L2 and NE.

PREVIOUS STUDIES

The previous studies on L2 speech production in Indonesia have scarcely been done and mainly based on auditory judgment and experience of teachers or researchers in teaching practice. Many scholars have described the characteristics of English pronunciation produced by Javanese learners. Nonetheless, studies about Javanese learners of English particularly in acoustical phonetics have rarely been conducted by Javanese or Indonesian linguists. In Indonesia, studies about non-native pronunciation are mainly based on phonological interpretation such as minimal pairs and listening judgment. There are only few researches related to the study.

Perwitasari (2015) writes about vowel durations in English as a second language among Javanese Learners. She finds that Javanese Learners of English seemed unaware of long and short duration of English vowels. Another Indonesian researcher, Widagsa (2015) found that Indonesian learners of English are strongly influenced by their mother tongue. Indonesian learners of English make a good impression only when the vowels in English are similar to Indonesian. It is proven by formant frequencies which are in common.

In Singapore, Deterding (2007) conducted a research related to speech production measuring Singaporean English (SE) Diphthongs. He outlined the distinctive characteristics among the Singaporean speakers of English compared with RP. He found that Singaporean speakers are less diphthongal than the corresponding vowels in standard British English. The use of a relatively monophthongal realization of the diphthongs by Singaporeans might be regarded as a distinctive characteristic of the local speech. Different results came from Tsukada (2008) who found that the production of English diphthongs by Thai learners of English were much closer to NE. Further, their production of the diphthongs might have been related to a large number of diphthongs in the Thai vowel inventory, which might have encouraged its speakers to substitute existing L1 categories for the English diphthongs rather than forming authentic new phonetic categories.

According to the previous explanation, in Indonesia, research related to acoustical phonetics especially in diphthongs of English as a second language have rarely been done. English diphthong of Javanese learners is an interesting topic to study for researchers. Thus, it can be said that this research is relatively new and needs to be conducted. This research will give novel contribution to ELT and other second language learning in Indonesia particularly in local language interference such as Javanese.

ENGLISH CLOSING DIPHTHONGS

Most varieties of English have several diphthongs. The most obvious diphthongs are the vowels of choice, mouth and price in most standard varieties of English. These diphthongs start with open vowels and then rise to close vowels, gener-

ally in the area of [i] or [u]. These are called closing diphthongs for this reason. Diphthongs are transcribed by the start and end points. For example, the vowel of choice is transcribed in RP as [Ti]: it starts with [T] and ends with [i]. Ogden (2009, pp.70-71) reports eight diphthongs in British English RP and divides in two main types; RP closing diphthongs and RP centering diphthongs.

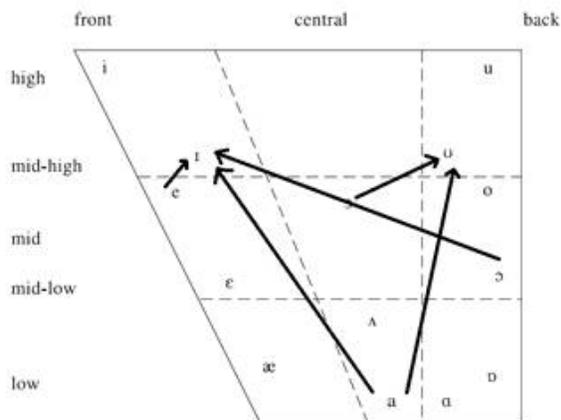


FIGURE 1: CLOSING DIPHTHONGS

JAVANESE DIPHTHONGS

Javanese, one of the most widely spoken languages in Indonesia, is a mother tongue of more than 75 million people (Crystal, 1997; Edi, et al., 2007). It is spoken mainly in Central Java and East Java. Javanese includes 8 vowel phonemes: 6 phonemes and 2 additional allophonic pairs [e] - [â] and [o] - [ɔ]. More recent studies confirm that Javanese vowels are grouped into 6 phonemes, including 4 allophonic pairs [i] - [I], [u] - [Ū], [e] - [â], and [o] - [ɔ]. The allophones of each vowel frequently occur in closed syllables. The standard Javanese of central Java is typically characterized as having six vowel phonemes. In the six vowels system, which appears to characterize the speech of our consultants, [â] and [ɔ] are in complementary distribution with [e] and [o], respectively.

Related to diphthongal vowels, Marsono (1999, pp. 54-58) claimed that Javanese has at least 5 diphthongs, one rising diphthong [ui] as in the words *uijo* 'extremely green', *cuilik* 'very small', *uireng* 'totally black' and the others are falling diphthongs.

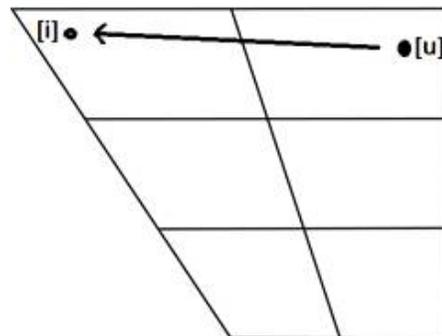


FIGURE 2: RISING DIPHTHONG IN JAVANESE

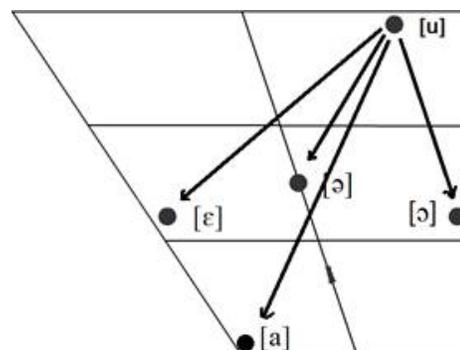


FIGURE 3: FALLING DIPHTHONGS IN JAVANESE

Marsono mentioned four falling diphthongs, they are:

- 1) Falling low open front [ua] as in *muarem* 'very satisfied', *uadoh* 'far away', *uanteng* 'very calm'.
- 2) Falling low open front [uɪ] as in *ngueyel* 'stubborn'
- 3) Falling low open back [uɔ] as in *luara* 'really painful' *duawa* 'very long'
- (4) Falling low open central [uɤ] as in *guedhe* 'enormous' *luemu* 'very fat'

However, these diphthongs only exist in east Java

and some of north east part of Central Java. The diphthongization in east Java is only used to exaggerate something.

METHOD

SUBJECTS

This experimental research involves five female subjects who were native speakers of Javanese aged 17-23. They used mainly Javanese in daily communication and lived in Yogyakarta. In addition, they had never been to English speaking countries. The subjects were all students of English Department of Universitas PGRI Yogyakarta (PBI-UPY) and all subjects had been studying English for at least 3 years. To ease the training process, the subjects were informed that the recording of their voice would be used for educational and experimental purposes. The audio data of English native speakers were taken from text-to-speech softwares.

STIMULI

All subjects produced a set of target words. The target words comprised eight diphthongal words such as *pay, five, home, now, join, near, hair, pure* which were inserted in a carrier sentence "*I say (diphthongal words) again*". During the recording, subjects repeated the sentence twice. The particular fillers or lexical sets were chosen to distract speakers' attention from the experimental words in order to encourage natural pronunciation.

RECORDING PROCEDURE

Before the recording process began, first, each subject completed a consent form, a brief questionnaire, which provided information about the subjects' native language and second language background. Second, subjects received a short introduction monologue which contained words

simulated for the recording. In order to get accustomed to the target stimuli, they were given time to read and practice.

Subjects were familiarized with the experiment and the procedures of recording. Each subject took a seat in front of a computer screen with active mode recording tools (audio recorders, and microphone). Once the stimuli appeared on the screen, subjects started to produce the sentences. The stimuli are presented in random order. Speech production were recorded in a sound-attenuated room and stored on a computer. Audio recording is treated confidentially and used for acoustic analysis. Last, the recorded speech sounds were annotated and segmented.

The recordings were made with ASUS X200MA notebook and a microphone placed at the distance of about 10 cm from the subjects' mouths. The subjects were recorded one by one in order to ease the analysis. This process took place in the language laboratory of Universitas PGRI Yogyakarta.

ACOUSTIC MEASUREMENT

Closing diphthongs are sounds which the vowel quality changes from relatively open position to a more closing position during the course of the vowel (Deterding, 2007). As the vowel quality is becoming less open, the F1 at the start point is expected to be different from the end point. The decrease of F1 would be an indication of how diphthongal the sounds are.

The recordings were analyzed using PRAAT 5.3.51. The software allows some features such as tracing the formant frequencies, choosing time point, and drawing waveform and spectrogram display.

TABLE 1: THE FORMANT FREQUENCIES AND DURATION OF L2

SYMBOL	MEAN				DURATION (SECOND)
	F1 START (HZ)	F1 END (HZ)	F2 START (HZ)	F2 END (HZ)	
eɪ	621,8	552,8	1374,4	2331	0,32394
ʌɪ	838,4	592,8	1138,8	1483,2	0,26866
əʊ	598	577,4	1119,2	987,6	0,3082
ɑʊ	793,8	539,2	1469,6	1093,4	0,35274
ɔɪ	600	490,6	1748	1946,4	0,3926

TABLE 2: THE FORMANT FREQUENCIES AND DURATION OF L1

SYMBOL	MEAN				DURATION (SECOND)
	F1 START (HZ)	F1 END (HZ)	F2 START (HZ)	F2 END (HZ)	
eɪ	542,6	421	2228,2	2318	0,22104
ʌɪ	890,4	802	1500,6	2006,4	0,28364
əʊ	629,6	445,4	1284,8	1175,6	0,2707
ɑʊ	663,6	626,6	1597,8	1222	0,2567
ɔɪ	403	503,2	1536,8	2305	0,27942

STATISTICAL ANALYSIS

To measure a difference between the production of L2 diphthongs and L1 English, this research conducted statistical analysis. An independent t-test for the groups was applied to test whether frequencies and speech duration was significantly different between groups.

RESULTS AND DISCUSSION

FORMANT FREQUENCIES

Diphthongs are the monosyllabic vowels which have two discernibly different points, one at the start and one at the end. Therefore, the soundwave data from the respondents are carefully analysed by defining the two target points. The first target is the starting point (vowel start) and the second is the ending point (vowel end). In measuring diphthongs, the vowel quality of vowel starts and vowel ends must be measured then both vowel qualities

will be plotted in a vowel chart in order to figure out the movement of the vowels.

The complete result of the the acoustic measurement of L2 and L1 English closing diphthongs is shown in the Appendix. Table 1 describes the mean of formant frequencies of English closing diphthongs and duration by Javanese Learners and Table 2 shows the mean of formant frequencies and duration of L1 English closing diphthongs

Based on the start points, several of the English closing diphthongs produced by Javanese learners of English were nearly identical to NE in regard to vowel height. To ease the comparison, the numerical data are plotted to bar chart.

The following bar chart illustrates the comparison of F1 value of English L2 and NE. Figure 4 demonstrates the start points of F1 while Figure 5 shows the end point fo F1.

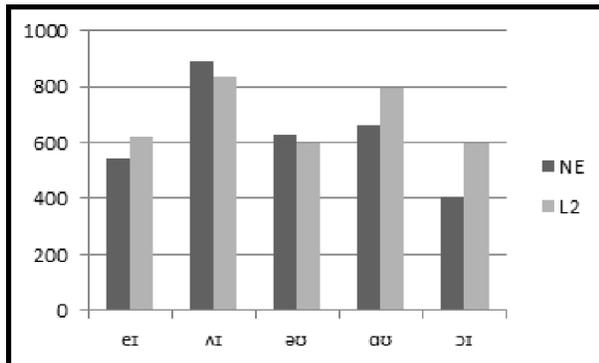


FIGURE 4: COMPARISON OF F1 STARTS OF NE AND ENGLISH L2.

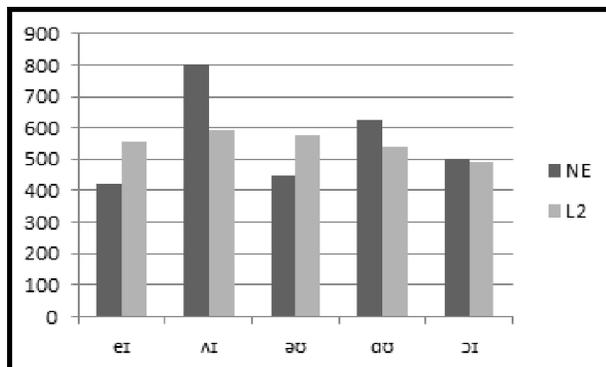


FIGURE 5: COMPARISON OF F1 ENDS OF NE AND ENGLISH L2.

It can easily be noticed from the previous tables and figures that the most similar result emerged in the pronunciation of /ɔɪ/ (F1 start NE 629,6 Hz and F1 start English L2 598 Hz). On the other hand, the most distinctive diphthong produced by English L2 was /ɔɪ/ (F1 start NE 403 Hz and F1 start English L2 600 Hz).

Unlike the start points, the average F1 value of diphthong /ɔɪ/ was close to the NE production (F1 end English L2 490,6 Hz and F1 end NE 503,2 Hz). Even so, the most disparate F1 end value was /æɪ/ with 592,8 Hz for English L2 and 802 Hz for NE.

To observe the decrease and increase of F1 value of English L2, the changes of F1 are presented in the following chart.

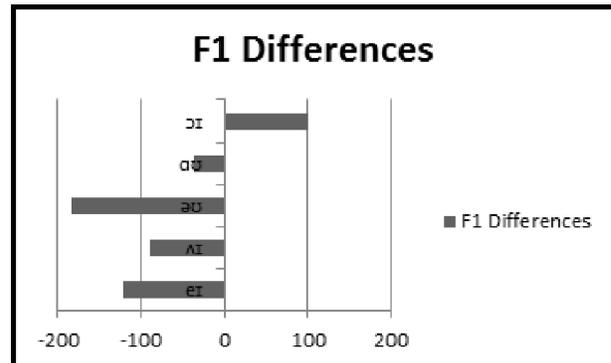


FIGURE 6: F1 CHANGES OF ENGLISH L2

Figure 6 demonstrates the decrease and increase of F1 value in closing diphthongs. The decreasing values are shown by the left bars. The right bar indicates the increase of F1 value. It can easily be spotted that there is a right bar for the sound /ɔɪ/ which indicates the significant difference among the diphthongal productions. The F1 value in the sound /ɔɪ/ surprisingly increased instead of decreasing. It simply puts that English L2 tends to widely open the mouth at the end of the diphthong when they pronounced the sound /ɔɪ/.

STATISTICAL RESULT

Closing diphthongs in English are /ej/, /æɪ/, /ɔɪ/, /əʊ/, /ɔɪ/ and expected to have decrease in F1. In general, the English closing diphthongs produced by English L2 were pronounced near NE in regard to the vowel height. In addition, the statistical analysis also confirms similar result. The following table shows the statistical analysis employed in this experiment.

Table 2. indicates that there is only one diphthong produced by L2 which is statistically different from L1. The result of the measurement of /ɔɪ/ is higher ($t(8)=1,995$, [mean= -254,6], [SD=212,232]) than t table which is 1.86.

TABLE 3: F1 CHANGE OF L1 AND L2

WORDS	GROUP	MEAN	SD	Df	T
eɪ	L2	-69	125,3	8	0,878
	L1	-121,6	47,56		
aɪ	L2	-245,6	87,331	8	1,212
	L1	-88,4	276,517		
əʊ	L2	-20,6	202,311	8	1,163
	L1	-184,2	240,86		
aʊ	L2	-254,6	120,13	8	1,995
	L1	-37	212,232		
ɔɪ	L2	-109,4	76,4087	8	-2,501
	L1	100,2	171,111		

*t label = 1.86

VOWEL CHART

In this case, since quadrilateral vowel chart requires the frequency of both formants (F1 and F2), the F2 is used to plot the numerical data Bark Scale. Thus, the numerical data are then plotted in a vowel chart (Bark Scale) to figure out how diphthongal they are. The following figure is the Bark Scale of average formant changes by respondents in producing English closing diphthongs.

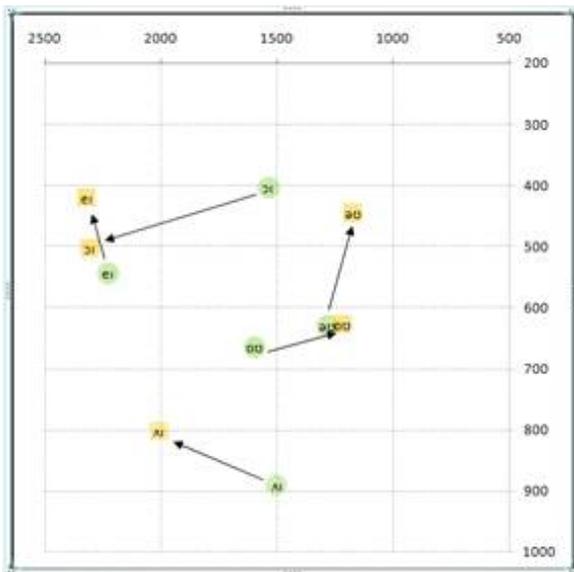


FIGURE 7: THE AVERAGE OF L2 FORMANT CHANGES OF EACH ENGLISH CLOSING DIPHTHONG IN BARK SCALE.

The greendots indicate the starting points of diphthongs and the yellow square dots represent

the ending points of the diphthongs. It can clearly be seen that there are variations among the diphthongal sounds. It can clearly be seen in Figure 5 that the diphthong /Tj/ was produced differently. The arrow points at down left side of the graph. It means that the end point of the diphthong was pronounced as low vowel with half open mouth.

TABLE 3 STATISTICAL MEASUREMENT OF DIPHTHONGAL DURATION

Words	Group	Mean	Sd	df	t
eɪ	L2	0,323	0,082	8	2,468
	L1	0,211	0,041		
aɪ	L2	0,268	0,049	8	0,305
	L1	0,283	0,098		
əʊ	L2	0,308	0,101	8	0,735
	L1	0,27	0,052		
aʊ	L2	0,352	0,06	8	2,553
	L1	0,256	0,058		
ɔɪ	L2	0,392	0,069	8	3,51
	L1	0,279	0,018		

*t label = 1.86

DURATION

In this research, the duration was only measured in the diphthongal words. The statistical analysis of L1 and L2 duration in producing diphthongal words are served in the following table.

Table 2. shows that there are significant difference in L2 duration. The differences are in the /ej/, /QŠ/ and /Tj/. It clearly indicates that L2 has longer duration in pronouncing diphthongal words than L1. The following figure demonstrates the difference.

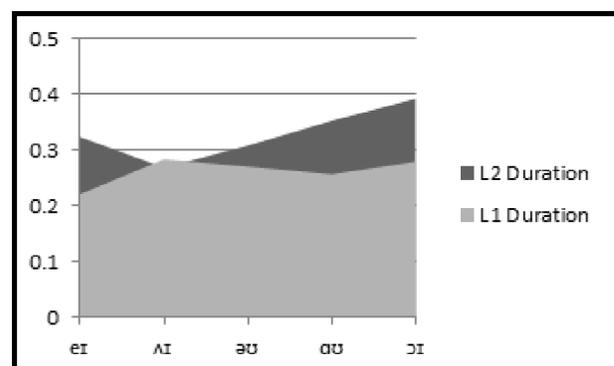


Figure 8. Duration of L1 and L2

Figure 9 shows that L2 tend to lengthen the duration particularly in /ej/, /QŠ/ and /Tj/. However, there was only one diphthong which was pronounced shorter than native speakers of English; the diphthong /æj/.

CONCLUSION

Javanese Learners of English do not have references to the set of diphthongs attributed to their first language phonetic features. Javanese are predicted to produce English diphthongs differently compared to native speakers of English. There are several findings which can be drawn based on the results. Admittedly, English closing diphthongs produced by L2 were not statistically different from L1, the different was only on the production of /QŠ/. It should be noted that there is a difference in the change of F1 values of the sound/Tj/. The differences answer the research question that the production of English diphthong by Javanese learners of English was different from the corresponding sounds produced by English native speakers. The other answered question was duration. It shows that there are significant differences in diphthongal duration produced by Javanese learners of English. They tend to lengthen the diphthongal words, three main findings stand out.

To put differently, the results of this measurement particularly in English closing diphthongs and the duration, support the hypothesis' prediction that the L2 failure in producing L2 vowels is related to the absence of diphthongal sound as a contrastive feature in L1. Javanese does not include diphthongs in its phonological system. It proves that the absence of those features in L2 can be a barrier for learners to study the language. The measurement results highlight a potential learning prob-

lem for Javanese learners of English. They are heavily influenced by their mother tongue.

This research is only a preliminary study which still requires further research related to acoustic features of English vowels and consonants produced by non-nativespeakers particularly Indonesian speakers or Indonesian regional language speakers. In this paper, I only report data of male respondents. Age is known to be an important factor of phonetic variation, aged-related would also be interesting topic for future research. Hopefully, this research would be worthwhile for language teaching in Indonesia.

ACKNOWLEDGEMENT

The author would like to thank the Directorate General of Higher Education, Indonesian Ministry of Research, Technology and Higher Education through Kopertis V, who in partnership with Universitas PGRI Yogyakarta funded this article. Any of the opinions, findings, statements and conclusions expressed in this article are solely the responsibilities of the author and do not necessarily expose those of the Kopertis V.

REFERENCES

- Baker, A & Goldstein, S. (2001). *Pronunciation Pairs: An Introduction to Sound of English*. Cambridge: Cambridge University Press.
- Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics*. Malden: Blackwell Publishing.
- Crystal, D. (1997). *The Cambridge of Encyclopedia of Language*. Cambridge: Cambridge University Press.
- Deterding, D. & Hvitfeldt, R. (1994). The Features of Singapore English Pronunciation/: Implications for Teachers, 15(1).

- Deterding, D. (2007). "The Vowels of the Different Ethnic Groups in Singapore" in *English in Southeast Asia: Literacies, Literatures and Varieties*. David Prescott, Andy Kirkpatrick, Isabel Martin and Azirah Hashim (eds.), Newcastle, UK: Cambridge Scholars Press, 2007, pp. 2–29.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press
- Hawkins, S., & Midgley, (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association Journal of the International Phonetic Association*, 35(35), 183–199. <http://doi.org/10.1017/S0025100305002124>
- Kreidler, Charles W. (2004). *The Pronunciation of English: A Course Book. Second Edition*. Victoria: Blackwell Publishing.
- Ladefoged, P. (2001). *A Course in Phonetics, 4th Edn*. Texas: Harcourt.
- Lado, R. & Fries, C.C. (1957). *English Pronunciation*. Ann Arbor: The University of Michigan Press.
- Ladefoged, P. (2001). *A Course in Phonetics, 4th Edn*. Texas: Harcourt.
- Lodge, K. (2009). *A Critical Introduction to Phonetics*. New York: Continuum International Publishing Group.
- Marsono. (1999). *Fonetik*. Yogyakarta: GadjahMada University Press
- Odgen, R. (2009). *An Introduction to English Phonetics*. Edinburgh University Press.
- Perwitasari, A., Klamer, M., Witteman, J., Schiller, N. O. (2015). Vowel duration in English as a second language among Javanese learners. In *The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow. Doi:<http://dx.doi.org/10.13140/RG.2.1.1752.5520>.
- Roach, P. (2009). *English Phonetics and Phonology: A practical course*. Cambridge: Cambridge University Press.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209-231.
- Subroto, E., Dwihardjo, M., and Setiawan, B. (2007). "Model Pelestarian dan Pengembangan Kemampuan Berbahasa Jawa Krama di Kalangan Generasi Muda Wilayah Surakarta dan sekitarnya". Laporan hasil Penelitian Tim Pascasarjana.
- Tsukada, K. (2008). John Benjamins Publishing Company. <http://doi.org/10.1075/eww.29.2.05tsu>
- Uhlenbeck, E.M. (1963). *Review of Beginning Javanese*, by Eleanor C. Home. *Lingua*, 12:69-86
- Widagsa, R. (2015). Acoustic Analysis of English Vowel Formant Frequencies Produced by Indonesian Native Speakers. Proceedings of The 2nd International Language and Language Teaching Conference. Yogyakarta: Universitas Sanata Dharma. 625-633.