The acquisition of English vowels by Javanese and Sundanese native speakers

Perwitasari, A.

5
General Conclusion

5.1 Introduction

The purpose of this research was to investigate the perception and production of English vowels by Javanese and Sundanese speakers (henceforth L2 learners). This chapter reports the general conclusions with respect to L2 speech learning by the Javanese and Sundanese speakers and implications for the teaching of English as an L2 teaching in Indonesia. It also presents the limitations and provides recommendations related to L2 learning.

The use of two or more languages in a community of speakers, i.e. diglossia, is common in Indonesia because the country is home to more than 700 local languages spoken by a total of around 255 million people. Many studies have carried out cross-linguistic comparisons between English as a foreign language and western European languages such as Spanish, German, Dutch, Swedish, Danish, and Italian. However, few studies have investigated Austronesian languages, such as Javanese and Sundanese - two widely spoken local languages in Indonesia. Since English has officially become a foreign language to study in Indonesian schools, most Javanese and Sundanese learners of English face some problems learning English sounds. Given the status of English as a foreign language in the multilingual context of Indonesia it is important to study the production and perception of the Javanese and Sundanese L2 learners. The empirical results of this study can be used to improve strategies to learn English as an L2.

This thesis has addressed the issue of Javanese and Sundanese speakers' perception and production of English sounds. The aim of the current study was to examine how cross-linguistic differences in the vowel systems of Javanese and Sundanese affect the perception and production of English sounds.

An often-discussed model, which is rooted in the production of second language sounds, is Flege's (1995, 1999, 2002) Speech Learning Model (SLM). According to SLM, L2 learners can accurately produce L2 sounds if they have accurate understandings on phonetic distance between L1 and L2 sounds. As a heuristic, within SLM, similar sounds are L2 sounds which are phonetically transcribed with the same base symbol as a sound in the learner's L1, but with at least one difference in diacritics (denoting length, nasalisation, aspiration or some other subtle phonetic difference). New sounds are L2 sounds which have no counterpart in the learner's L1 that is written the same base symbol. Similar sounds are predicted to block the formation of a new phonetic category formation in L2 learning process. Instead, the existing L1

category is extended over the years to include also the L2 tokens, with a shift of the prototype roughly midway between the original prototypes of the L1 and L2, so that after completion of the acquisition process the category is wrong both for the L1 and the L2 – without the learner being aware of this. New L2 sounds, on the other hand, enable the learners over the years to develop new L2 categories, which are authentic in the L2 and do not interfere with the categories in the learner's L1 (Flege, 1995).

Another model on L2 sound learning is Escudero's (2005) Second Language Linguistic Perception model (L2LP). L2LP posits that the acquisition of L2 sounds should match the acoustic properties of L1 sounds (Escudero & Boersma, 2004; Escudero, 2005; Escudero & Williams, 2012). According to L2LP, the L2 sounds to be acquired are categorized into one of three scenarios, i.e. *new*, *subset* and *similar*. The *new* scenario occurs if the L1 perception grammar permits fewer perceptual categories than are required for native perception of the L2. As a result, the L2 environment produces phonological differences that do not exist in the L1 (Escudero, 2005).[4] The subset scenario occurs if the L1 perception grammar outputs more categories than the perception of the L2 requires. Thus, the L2 categories constitute a subset of L1 categories.[5] In the similar scenario, the L1 perception grammar outputs the same number of categories as the target sounds require, because the L1 and L2 categories are phonologically equivalent. This could be called the no (or neutral) transfer scenario.

In Chapter 1, we have introduced two L2 perception models, i.e. SLM and L2LP. SLM predicts that similar L2 sounds (with respect to L1) will always remain relatively difficult to produce and perceive. New L2 sounds, however, will ultimately cease to be perception and/or production problems because, given sufficient time/exposure, a new sound will trigger the formation of a new phonetic category in the L2 without any L1 interference. L2LP predicts that the acquisitions of the new L2 sounds are relatively difficult, and that new sounds are a greater source of difficulty than similar sounds, during the time that the L2 acquisition is in progress. The present thesis tested which of these two models was best supported by the results of the acquisition of the vowels by Javanese and Sundanese L2 learners of English.

We also tested the L2 Feature-dependent hypothesis in L2 speech production, which states that if duration is not used to signal

[4] This scenario is called underdifferentiation of the L1 relative to the L2 in Lado's (1957)'s transfer theory.

[5] This is equivalent to the overdifferentiation scenario in the transfer theory.

phonological contrasts in the learner's L1, it will be difficult to produce (McAllister et al., 2002). We compared this hypothesis with the Desensitization Hypothesis (Bohn, 1995). The Desensitization Hypothesis states that, even if the duration feature is not used contrastively in the L1, it will not be difficult to produce because duration cues are acoustically salient and easily acquired. The present thesis tested whether the Feature-dependent hypothesis vs. Desensitization Hypothesis was supported by our results.

5.2 L1 production of Javanese and Sundanese Speakers

The study in Chapter 2 set out to extend those of Van Zanten and Van Heuven (1984) and Van Zanten (1986), investigating to what extent the Standard Indonesian vowels spoken with a Javanese or Sundanese accent are similar to the Javanese and Sundanese vowels produced by the Javanese and Sundanese speakers.

The results showed that the Javanese female and male speakers produce a high front /i/, a mid front /e/, a low /a/, a mid central /ə/, a high back /u/, and a mid-back /o/. The Javanese schwa /ə/ is considerably higher than that of the Sundanese speakers. The Javanese produced schwa remarkably closed and front.

The results also confirmed that Sundanese female and male speakers produce a front /i/, mid front /e/, low /a/, high central /ɨ/, mid central /ə/, high back /u/, and mid-back /o/. Vowel /ɨ/ is produced in closed central position, only for the male speakers. The Sundanese female speakers produced /ɨ/ with an unclear distinction with the production of schwa.

Based on the descriptive analyses in Chapter 2, it was apparent that the Javanese schwa /ə/ is remarkably close and front. This might have resulted in a problem with pronouncing English words such as *cup, butter* and *but*. Additionally, /a/ appears to be produced in the front part of the mouth by the L2 learners, possibly because their L1 /a/ was relatively (as compared to English) front and open. This thesis discusses these potential pronunciation problems with schwa /ə/ and /a/ in chapter 4.

5.3 L2 Perception of Javanese and Sundanese Speakers

Chapter 3 of the thesis examined how Javanese and Sundanese speakers acquire perception of English (L2) vowels. One aspect analysed in this study was whether the sound category (similar vs. new as compared to

the L1 vowel inventory) affects the L2 sound perception by Javanese and Sundanese listeners. Another aspect analysed was whether the phonetic distance between L2 target and distractor influence the L2 learners' task performance. To find out the answers, we carried out a mouse tracking experiment by specifically measuring the Error rate, Area Under the Curve (AUC), initiation time, reaction time and velocity profiles of the participants.

First, we examined whether the familiarity with the L2 sounds (new vs similar) influence the L2 learning. The results of Error rate and AUC showed that listeners had more problems when they perceived new than similar sounds. The L2 learners made more errors when the target L2 vowel was spectrally close to the competitor. Hence, the results support the L2LP model indicating that perceiving new sounds create more problems to L2 learners than similar sounds.

Interestingly, the results of velocity profiles did not produce any clear evidence for either of the L2 learning models regarding the effect of L2 sound familiarity. In the early time window (226 - 750 ms post-stimulus), the Javanese learners showed a trend of high velocity profiles regardless the sound category, while the Sundanese learners showed the higher velocity profiles than the American listeners for the new sounds. In the later time window (751 - 1125 ms post-stimulus), the L2 learners moved their mouse pointer more slowly than the American L1 listeners, for both new and similar target sounds. Altogether, the results of this new behavioral measure could not clearly identify specific perception problems by the L2 learners.

Second, we investigated whether phonetic distance affects the L2 sound leaning. We assume that the L2 learners will find it hard to perceive L2 sounds as the phonetic distance is spectrally close. The results of error rates and AUC showed that the L2 learners showed relatively high mistakes when the acoustic distance between target vowel and competitor is small. Thus, the results confirm that the smaller the acoustic distance is, the hard it is for the L2 learners to discriminate the L2 sounds.

To summarize, our data are in line with the L2LP model suggesting that L2 vowels which are acoustically new in the Javanese and Sundanese native languages may be harder to perceive than similar vowels as the former require learners to create new categories and perceptual mappings and to integrate them with the already categorized sounds.

5.4 L2 production of Javanese and Sundanese Speakers

In this section, we discuss the production of L2 English vowels by Javanese and Sundanese speakers. Models such as SLM and L2LP suggest that vowels which are difficult to perceive are also hard to produce. SLM postulates that similar sounds between the learner's native language L1 and the target language (L2) are more prone to trigger learning difficulties than new sounds because L2 learners will ultimately establish a new phonetic category which is different from any L1 sounds rather than considering it an equivalent of some L1 sound. SLM predicts that the L2 learners may continue to experience difficulties with the similar vowels /iː/ and /uː/ but would learn to be capable of successfully producing the new vowels /ɪ, ɛ, ʊ, æː, ɑː, ɔː, ʌ, and ɝ. On the other hand, L2LP holds that the production of new L2 sounds is more difficult than producing similar L2 sounds. Thus, the Javanese and Sundanese speakers may struggle more with producing the new L2 vowels than the similar L2 vowels.

The F_1 values of the new L2 vowels /ɑː/ and /ɪ/ were considerably lowered by the L2 learners as compared to native speakers. In addition, the F_2 values of the new L2 vowel /ɪ/ were produced significantly higher by the Sundanese speakers as compared to the American speakers. The lowered F_1 values indicate that the L2 learners produced vowels /ɑː/ and /ɪ/ with a more raised tongue. The high F_2 values of vowel /ɪ/ indicate that it is produced more frontally by the Sundanese speakers than by the American speakers. The Javanese speakers did not show any significant differences with the production of English new and similar vowels in F_2 values. Since the new vowels proved to be relatively difficult, these results, at least at first sight, would support L2LP rather than SLM. Again, however, the Indonesian learners of English who served as the participants in this dissertation may still be in the intermediate stage of the L2 acquisition process, so that the possibility that the new sound would ultimately, after the completion of the L2 acquisition, be set up as an authentic category in English (i.e., indistinguishable from the L1 counterpart even as judged by native listeners of the target language) cannot be ruled out.

5.5 Overall Conclusion Regarding L2 Learning Models

Second language learning models such as Speech Learning Model (SLM, Flege, 1987, 1995, 2002, 2003) and the Second Language Linguistic Perception model (L2LP, Escudero, 2005, 2006, 2009) explain L2

acquisition problems based on the similarity of L1 and L2 sounds. According to SLM, L2 sounds, which are similar to L1 sounds, will ultimately remain more difficult than new sounds. L2LP, on the other hand, predicts that new L2 sounds will be more difficult than similar sounds throughout the L2 acquisition process.

Chapter 3 explains that the L2 learners particularly show difficulties perceiving L2 sounds when the target vowel was spectrally close to its L2 competitor vowel. The results, however, additionally showed that perception differences between the native and non-native speakers were largest for new vowels. Chapter 4 indicated that L2 learners have difficulty in producing the similar vowels /iː/ (not for vowel /uː/), but also with new vowels (for Javanese: /ɪ, ɛ, æː, ɜ˞, ʌ, ɔː, ɑː, ʊ/; for Sundanese: /ɪ, ɛ, æː, ɜ˞, ʌ, ɑː, ɔː/). Taken together, the results of L2 perception give partial support to both models, but more to L2LP than SLM.

The next question would be if and how the L2 learning models can be supported simultaneously. One possibility is that the L2 learning models focus on different stages of learning. SLM focuses on *experienced* learners. L2LP, in contrast, assesses relatively unexperienced learners, which match with the speakers in the current study. Therefore, relative difficulty with *new* vowels, as observed in the present thesis, was to be expected.

We additionally tested two feature specific hypotheses, i.e. Feature-dependent Hypothesis (FH) by McAllister, Flege, and Piske (2002) and the Desensitization Hypothesis (DH) by Bohn (1995). FH predicts that L2 learners in the present work will have difficulties acquiring duration cues because these are not used contrastively in their L1. DH, in contrast, predicts that the L2 learners would have no dificulties acquiring the duration cues because duration remains relatively easy to access even if it is not used contrastively in the learners L1. In the present work, the L2 learners shortened all target vowels, both long/tense and short/lax. Specifically, they over-shortened short vowels. From this perspective, the results seem to support FH indicating that L2 learners have difficulties in producing duration in a native-like manner because the durational information is not part of their L1 phonology. On the other hand, the results showed that the Indonesian learners produced the relative durations in an almost native-like manner. The L2 durations correlated at better than $r = .800$ with the American L1 durations, and the long and short vowel categories were quite clearly separated. This would force us to accept DH at least in part. Moreover, there are alternative explanations for the

overall (apparent) shortening of the L2 vowels by the Indonesian learners. The American speakers may have used longer (vowel) durations than is normal, since they may have been developed a strategy to speak more slowly to Indonesian listeners in order to be better understood. A slow rate of delivery is listed as a prominent characteristic of this so-called foreigner talk. Moreover, the Indonesian speakers will have pronounced the target vowels followed by a voiceless [t] (in stead of the L1 [d], and failed to apply vowel lengthening before the coda obstruent. The mere fact that the Indonesians adequately differentiated between short lax vowels and long(er) tense vowels is further, strong and positive evidence in favor of DH. There are (weak) indications that the Indonesians habitually speak with very short vowel durations, shorter even that the short vowels in English and Dutch. The vowel durations we measured for the six or seven vowels of Javanese are shorter that what is normally reported. Very short vowel durations were also found by Van Zanten and Van Heuven (1983) in their sample of ten speakers of Standard Indonesian, five of whom were of Javanese origin. On balance, then, our data lend support to both FH and DH, and - certainly in hindsight - it now seems that the two hypotheses are not mutually exclusive. Late L2 learners may remain more sensitive to duration than to other cues, and tune on to the distinctive length effects in the L2. But at the same they may have lost some (but not all) all sensitivity to duration when it is not part of their native-language phonology, and/or they may transfer the habit of producing very short vowel durations from their L1 to the L2.

5.6 Specific Relative Difficulties of the L2 Perception and Production of English Sounds

English vowel perception and production appear to be problematic for Javanese and Sundanese learners of English as the results of the experiments have shown.

The vowel pair /æ: - ɛ/ is the most difficult L2 contrast to be perceived by our Indonesian L2 learners of English. Table 5.1 summarizes the production problems faced by our Javanese and Sundanese participants' as compared to American native speakers.

Table 5.1 The summary of the English sounds which are difficult to be produced by the Javanese and Sundanese speakers.

L1	Target vowels	Mispronounced as
Javanese	/iː/	/ɪ/
	/æː/	/ɛ/
	/ɑː/	/ʌ/
	/ɑː/	/ɔː/
Sundanese	/æː/	/ɛ/
	/ɑː/	/ʌ/
	/ɑː/	/ɔː/
	/ʌ/	/æː/
	/ɑː/	/æː/

The results of vowel perception in this study reveal the following relative difficulties of the L2 learners (Chapter 3):

a. Generally, the Javanese and Sundanese learners have relative difficulty perceiving new vowels (relative to the L1 vowel inventory)
b. Phonetic distance affects the L2 sound learning, showing that the Javanese and Sundanese have more perception difficulty as the phonetic distance is closer.
c. The effect sizes of interactions showed that the phonetic distance between the L2 target and competitor vowel is almost an order of magnitude larger than that of sound category (new vs similar sounds).

Acoustic measurements of the English vowels reveal that Javanese and Sundanese learners of English have problems in producing English vowels (Chapters 4). The results reveal the following relative difficulties of the L2 learners:

a. Javanese and Sundanese learners of English produced lower F_1 values than the English speakers.
b. The English vowel durations of Javanese and Sundanese speakers were shortened for both long and short vowels.
c. The Javanese and Sundanese speakers have difficulty contrasting English vowel pairs whose members are adjacent in the articulatory vowel space using spectral parameters while the use of duration is used contrastively, possibly with over-shortening.

5.7 Implications for L2 Teaching

The following recommendations can be made based on the current study with regard to the learning of English as a second language by Javanese and Sundanese university students, and possibly by Indonesian students in general.

The findings confirm that Javanese and Sundanese speakers do not accurately perceive the new L2 vowels /ɑː, ʌ, æː, ɛ, ɪ, ʊ/ and the similar L2 vowels /iː, uː/. It is therefore recommended that English teachers should design vowel identification tasks to help students perceive the differences between all adjacent pairs of the above-mentioned vowels. Laboratory discrimination training on English vowels using the MouseTracker software or other computer-based online training programs, such as the Perception of Spoken English (POSE) test <https://posetest.com/>, may exert a positive influence on speech perception.

As the findings in the current study show that the Javanese and Sundanese speakers have difficulty contrasting intended vowels using spectral parameters, we suggest that the English teaching and learning process for Javanese and Sundanese speakers focus on increasing the F_1 for the vowels /ɑː/, /ɪ/, /æː/, and /iː/. Javanese and Sundanese speakers should be trained to produce these vowels with more openness and a frontal tongue position. Secondly, as we found that the duration of Javanese and Sundanese speakers was shortened for both long and short English vowels, we recommend that Javanese and Sundanese speakers receive phonetic training focusing on lengthening short and long vowels, or perhaps talk more slowly in English than they are used to in their native language. One possibility to achieve these goals would be to use software to let L2 learners produce target vowels, record formant frequencies and durations, and then provide real time feedback to improve pronunciation. Using this approach, we expect that English pronunciation can be significantly improved among Javanese and Sundanese learners of English.

5.8 Limitations

We would like to outline the limitations of this study.

First, the current study focuses on vowels alone, neglecting other features such as consonants or consonant clusters which might compromise the intelligibility of Indonesian speakers of English as much or even more than incorrect pronunciation of the vowels. Thus, we

suggest that further studies investigate how consonants or consonant clusters influence second language acquisition. However, it is important to note that the spectral and temporal properties of vowels play a role in speech intelligibility (Walker, 2001). Therefore, this study can be a useful starting point in addressing intelligibility problems in Javanese and Sundanese L2 speech learning.

The second limitation concerns the group of participants. The sample of the present study does not allow generalization to all native speakers of Javanese and Sundanese and all other Indonesian local languages because the characteristics of students from private universities, uneducated adults, and people from other regions may be different. L2 production and perception for other Indonesian local languages will also play a different role than it did in the case of Javanese and Sundanese. Ideally, we would have tested more groups from other Indonesian local languages. In the future, other researchers may include native speakers of other Indonesian local languages to illustrate a broader pattern of English vowel production and perception difficulties.

Third, the relatively small number of L1 participants tested, especially in the L1 production experiment, can be considered one of the limitations of this study. The number of Javanese (4) and Sundanese (4) native speakers in the L1 production test was very small. Testing more speakers of more languages will form a clearer description of the L1 vowel system of Indonesian local languages.

Finally, in this thesis we used a mouse-tracking technique to test L2 perception. It might have been harder for the non-native speakers of English to perform this task than for the native speakers of English. This might explain why the non-native speakers moved the mouse more slowly. It is interesting to see in the mouse-tracking data that the initial time window (226 - 750 ms post-stimulus) was not able to accurately show which specific English vowels are difficult to perceive. This may be due to the participants' lack of familiarity with the use of the computer mouse with which they interacted in the L2 perception experiment. This unfamiliarity may have influenced their mouse movements. We tried to control this issue by including practice trials in order to familiarize participants with the mouse-tracking software. Although we attempted to minimize familiarity effects in this way, we cannot exclude the possibility that the non-native speakers had a harder time to adjust to the task, resulting in non-linguistic differences between the native and non-native groups.

5.9 Future Research

The present dissertation and the studies reported in it cover only a small proportion of topics that may profitably be studied as a potential source of information on how the teaching of English pronunciation and of developing adequate listening comprehension in English could be improved for Indonesian learners of English as a foreign language.

The logical first step towards developing an insight into where the potential learning problems might reside would be to study in depth the perceptual assimilation patterns applied by Indonesian listeners when they are first confronted with the sounds of English. The Indonesian listeners' task would be to indicate for each sound of English (vowels and consonants) which sound in their native language they consider the foreign sound a token of, and how good a token of that native category it is. The results of this procedure will allow us to to assess, among other things, which English vowels map onto which Javanese or Sundanese vowels, and how well they fit the L1 vowel categories. This information, in turn, may be fruitfully used to rank-order the English vowels and vowel contrasts along a scale of difficulty for Indonesian learners of English. Examples of such studies are available in the literature and may serve as a blueprint for the kind of exercise required (e.g. Tsukada et al., 2005 for Korean learners of English; Sun & Van Heuven 2007 for Mandarin learners of English).

The present studies have not considered the intelligibility of Javanese and Sundanese accented English. The next step in the investigation should be to have groups of listeners, native as well as non-native listeners of varieties of English, identify the vowel sounds produced by our speakers. For lack of human native listeners, we have have taken recourse to computer-simulated (native) listeners using Linear Discriminant Analysis to generate a model that arguably approximated a human American native listener. This model can and should be verified using actual human listeners. It will be impossible to subject all 40 speakers (20 Javanese, 20 Sundanese) to such perceptual identification. Instead a smaller number of representative learners can be selected from the larger group, for instance using the individual scores obtained in the LDA as a criterion. Representative speakers would then be the persons closest to the centre of the distribution (see Wang, 2007; Wang & Van Heuven, 2014 as an example of how this can be done).

The perceptual awareness and production of the marked voiced – voiceless contrast in coda obstruents would be a topic that should be

studied with high priority. The perception and production of coda consonants, which are largely absent from the phonologies of Indonesian languages are expected to be a major challenge to Indonesian learners of English. It is important to find out whether the Indonesian speakers in the present study were able to produce a properly voiced word-final [d]-sound, accompanied by the lengthening of the preceding vowel, as is usual in native English. It would be worthwhile, for instance, to repeat the study reported in Chapter 4, replacing the target words by counterparts ending in /t/: e.g. *heat*, *hit*, *bet*, *hat*, *Baht* ('Thai currency unit'), *bought*, *put*, *hoot*, *hurt*, *hut*. The results of such a follow-up study would allow us, among other things, to assess whether indeed the difference in vowel duration between the Americans and the Indonesians would be reduced.

A separate project would be to map out the perceptual representation of the vowel system of English in the mind of Indonesian learners. American English has ten monophthongs, which differ in vowel quality and duration. An artificial vowel space can be defined by the acoustic parameters F_1 (10 perceptually equidistant steps, capturing vowel height), F_2 (10 equidistant steps, capturing constriction place and lip rounding) and duration (5 steps). This yields a vowel set of (less than) 500 types, which can be embedded in a foces carrier, and offered to L1 and L2 listeners for perceptual identification and judgment of typicality, following the examples of e.g. Van Zanten and Van Heuven (1984) for Indonesian vowels, Van Heuven (1986) for Dutch vowels or Van Heuven (2017) for English vowels. The results will probably show that vowel quality and duration determine the vowel categories in roughly equal proportion for the L1 listeners of English, but the duration will outweigh the spectral cues for the L2 listeners, in much the same way that was found by Van Heuven (1986) for Dutch vowels as perceived by L1 Dutch and L2 Turkish learners.

As explained in Chapter 1, more than 700 languages are spoken in the Indonesian archipelago. Most of these, including Javanese and Sundanese, belong to the Austronesian language family, and therefore share many structural properties. The research reported in the present thesis has shown that the two learner groups pronounce the English vowels in more or less the same way. This should not come as a surprise considering that the vowel systems of Javanese and Sundanese are virtually the same, and differ only in the presence of a high central vowel in Sundanese, which is absent from the inventory of Javanese and of most other Austronesian vowel systems. A notable exception is Batak, an indigenous Austronesian language spoken by over two million

speakers on the isle of Sumatra. Batak has a five-member vowel inventory without any central vowel (Van Zanten & Van Heuven, 1984), and has contrastive stress (Van Zanten & Van Heuven, 1997) instead of either fixed prefinal, variable or perhaps even no word stress at all, as is often claimed for other Austronesian languages (Van Zanten & Van Heuven, 1998, 2004; Goedemans, & Van Zanten 2007; Van Heuven, Roosman, & Van Zanten, 2008; Maskikit-Esed & Gussenhoven, 2016).

The presence of a second (high) central vowel in Sundanese would seem to offer no advantage to a learner of English, since English has no vowel in that part of the vowel-quality space that moght benefit from such a category. The absence of any central vowel, however, may be a source of learning difficulty for Batak learners of English. More generally, the question can be raised whether the teaching of English to speakers of different Austronesian languages should be different depending on the learner's specific L1, or whether a one-size-fits-all approach would be equally effective. As a first approximation to this issue, a study can be done to assess how well speakers of different Austronesian languages are able to determine whether an Indonesian speaker belongs to their own regional language community or not, and in the latter case, if they are able to identify the specific L1 – based on both the way they speak Standard Indonesian (which is a second language to most Indonesians) and how they pronounce English, along the lines sketched by Cui and Van Heuven (2011) for related languages spoken in China, or by Van Heuven and Gooskens (2017) for Scandinavian speakers of English. If the Indonesians are not able to differentiate their own pronunciation of the Standard language or of English from that of other regional groups of Indonesian speakers, there would be no point in developing English teaching materials for different regional learner groups in Indonesia.

The studies proposed here may serve to determine the causes of incorrect production and /or perception of the English vowels by Indonesian learners. They do not address the issue how the problems can be overcome. The pedagogy of the teaching of oral foreign-language skills is still very much in its infancy. As I said elsewhere in this thesis, computer-assisted teaching offers enormous advantages. Perceptual skills at all relevant linguistic levels (sound discrimination, word recognition, global listening comprehension) in the foreign language can be trained with computer feedback, and practically all aspects of foreign-language pronunciation, including the correct use of speech melody, can be practiced with computer-assisted supervision. Future research is needed to determine if the oral skills can be acquired more

quickly and effectively if the technological aids zoom in on specific perception and production problems that spring from interference phenomena between Indonesian and English.

References

Bohn, O. S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279-304). Baltimore, MD: York Press.

Cui, R., Heuven, V. J. van (2011). Mutual intelligibility of English vowels by Chinese dialect speakers. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, City University of Hong Kong, Hong Kong (pp. 544-547).

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551-585.

Escudero, P., & Williams, D. (2012). Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America, 131*, EL406–EL412. DOI: 10.1121/1.3701708.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. LOT dissertation series 113. Utrecht: LOT.

Escudero, P. (2006). *Second language phonology: The role of perception*. In M. Pennington (Ed.), *Phonology in context* (pp. 109-134). Houndmills, Basingstoke, Hampshire, New York: Palgrave Macmillan.

Escudero, P. (2006). The phonological and phonetic development of new vowel contrasts in Spanish learners of English. In. B. O. Baptista & M. A. Watkins (Eds.), *English with a Latin beat: Studies in Portuguese/Spanish-English interphonology* (Studies in Bilingualism 31, pp. 149-161). Amsterdam: John Benjamins.

Escudero, P. (2009). Linguistic perception of 'similar' L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 151-190). Berlin: Mouton de Gruyter.

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

Flege, J. E. (1995). Second language speech theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic*

experience: Theoretical and methodological issues (pp. 233-277). Baltimore: York Press.

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. (2002). Interactions between the native and second-language phonetic systems. In H. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development*, Papers in honor of Henning Wode (pp. 217-44). Trier: Wissenschaftlicher Verlag Trier.

Flege, J. E. (2003). Assesing constraints on second-language segmental production and perception. In Schiller, N. & A. Meyer (Eds.), Phonetics and phonology in language comprehension and production: Differences and similarities. Berlin: De Gruyter.

Goedemans, R. & Van Zanten, E. (2007). Stress and accent in Indonesian. In V. J. van Heuven & E. van Zanten (Eds.), *Prosody in Indonesian languages*. Utrecht: LOT. 35-62.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Maskikit-Essed, R., & Gussenhoven, G. (2016). No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay, *Phonology*, *33*, 353-389.

McAllister, R., Flege, J., & Piske, T. (2002). The Influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonia. *Journal of Phonetics, 30*, 229-258.

Sun, L., & Van Heuven, V. J. (2007). Perceptual assimilation of English vowels by Chinese listeners. Can native-language interference be predicted?, in B. Los, & M. van Koppen, (Eds.), *Linguistics in the Netherlands 2007* (pp. 150-161). Amsterdam: John Benjamins.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, *33*: 263-290.

Van Heuven, V. J. (1986). Some acoustic characteristics and perceptual consequences of foreign accent in Dutch spoken by Turkish immigrant workers. In J. van Oosten, J. F. Snapper (Eds.). *Dutch Linguistics at Berkeley, papers presented at the Dutch Linguistics Colloquium held at the University of California, Berkeley on November 9th, 1985* (pp. 67 – 845). Berkeley: The Dutch Studies Program, U.C. Berkeley.

Van Heuven, V. J. (2017). Perception of English and Dutch checked vowels by early and late bilinguals. Towards a new measure of language dominance. In: Simone E. Pfenninger & Judit Navracsics (eds.) *Future research directions for Applied Linguistics* (Second Language Acquisition 109). Bristol, Buffalo, Toronto: Multilingual Matters (pp. 73-98).

Van Heuven, V. J. & Gooskens, C. (2017). An acoustic analysis of English vowels produced by speakers of seven different native-language backgrounds. In Wieling M., Kroon M., Noort G. van, Bouma G. (Eds.). *From Semantics to dialectometry. Festschrift in honour of John Nerbonne,* no.32 London: College Publications (pp. 137-147).

Van Heuven, V. J., & Van Zanten, E. (1997). Effects of substrate language on the localization and perceptual evaluation of pitch movements in Indonesian. In *Proceedings of the 7th International Conference on Austronesian Linguistics*, Rodopi, Amsterdam/Atlanta, (pp. 63-80).

Van Heuven, V. J., Roosman, L.M., & Van Zanten. (2008). Betawi Malay word prosody. *Lingua, an International Review of General Linguistics, 118*(9), 1271-1287.

Van Zanten, E. (1986). Allophonic variation in the production of Indonesian vowels. *Bijdragen tot de Taal-, Land- en Volkenkunde, 142*, 427-446.

Van Zanten, E., & Van Heuven, V. J. (2004). Word stress in Indonesian: fixed or free? *NUSA, Linguistic Studies on Indonesian and other Languages in Indonesia, 48*, 1-18.

Van Zanten, E., & Van Heuven, V. J. (1983). A phonetic analysis of the Indonesian vowel system, a preliminary acoustic study. *NUSA, Linguistic Studies of Indonesian and other Languages in Indonesia*, *15*, 70-80.

Van Zanten, E. & Van Heuven, V. J. (1984) The Indonesian vowels as pronounced and perceived by Toba Batak, Sundanese and Javanese speakers. *Bijdragen tot de Taal-, Land- en Volkenkunde, 140*, 497-521.

Wang, H. (2007). *English as a Lingua Franca. Mutual intelligibility of Chinese, Dutch and American speakers of English.* LOT Dissertation series 147. Utrecht: LOT.

Walker, R. (2001). Pronunciation for international intelligibility. Karen's linguistics issues: Free resources for teacher and students of English. *English Teaching Professional Magazine, 2*.

Wang, H., & Van Heuven, V. J. (2014). Is a shared interlanguage beneficial? Mutual intelligibility of American, Dutch and Mandarin

speakers of English. In: Doel R. van den, Rupp L. (Eds.), *Pronunciation Matters. Accents of English in The Netherlands and elsewhere*. Amsterdam: VU University Press (pp. 175-194).