

The Acquisition Process of L2 Pronunciation —based on the acoustic analysis of English vowels produced by Japanese learners

Toru Isono

要 旨

本論文では、第二言語音声習得研究において長年議論されてきた、「どうしてある特定の音素が他のものよりも早く習得されるのか」という疑問に答えるため、日本人英語学習者から得た英語母音データを音響分析し、その結果を踏まえて第二言語音声習得プロセスのモデルを提案する。まず、中間言語という概念の説明をしてから、第一言語習得と第二言語習得のプロセスの違いを確認する。次に、英語母音の音響分析の概要を紹介し、その結果を考察する。そして最後に、本論文での全ての議論をもとに第二言語音声習得プロセスのモデルを提案し、その有効性を論じる。

キーワード： Second Language Acquisition (第二言語習得), pronunciation (発音), the acquisition model (習得モデル), acoustic phonetics (音響音声学), interlanguage (中間言語)

1. Introduction

Nowadays, the acquisition process of Second Language (L2) pronunciation is regarded as a creative process by which L2 learners produce an internalised representation of the regularities they find in the linguistic data to which they have been exposed through their interaction with their environment. To put it another way, L2 learners are supposed to be able to construct and develop their

own regular and systematic phonological systems as well as native children. The phonological systems that lie midway between a learner's First Language (L1) and the L2 are named *interlanguage phonology*, and the creative process is called *the continuum of interlanguage phonology*.

Concerning the nature of the continuum of interlanguage phonology, some pieces of research conducted in the 80's (e.g., Flege, 1980; Major, 1987a) clarified the following two points. One is that L1 interference is dominant in L2 phonological errors, especially at the beginning stage. The other is that the features of L1 interference are replaced by near-L2 features step by step after a certain period of learning. Based on these facts, we can probably say that the nature of the continuum of interlanguage phonology is a continuum that the origin is a learner's L1, and the features of the L1 sound system are gradually replaced with those of the L2 sound system. In other words, the acquisition process of L2 pronunciation is the selection process that decides which L1 features are transferred and which are not, and also which L2 features are taken into the systems of interlanguage phonology. The important thing is that it is still not clear why there are some L1 features more likely to be transferred than others, and why there are some L2 features more likely to be acquired than others.

In this paper, by examining the acquisition process of English vowels by Japanese learners, we will clarify a feature which affects the decision of the above L2 learners' selection process, and propose the model of the acquisition process of L2 pronunciation.

2. Kiparsky and Menn's Framework for First Language Acquisition in Pronunciation

One of the basic concepts of interlanguage is that the process of interlanguage is similar to that of First Language Acquisition (FLA) in some points. Therefore, we start the discussion by referring to the work which proposes a framework of acquisition of phonetic repertoire by native children.

The cognitive approach triggered by Chomsky claims that children do not only imitate the input they receive, but they have the ability of establishing their own hypotheses against the input and also the ability to test these hypotheses. Although this idea was developed in the domains of grammatical and morpheme studies, it was also accepted in the acquisition of pronunciation. For example, Kiparsky and Menn (1987) claim that the process of pronunciation acquisition by native children is a 'problem-solving' activity from the earliest stages, and propose a framework consisting of three stages, which is presented on the next page, for the acquisition model of L1 sounds.

Stage A is constituted by the children's hypotheses about the underlying representations of the adult language they are acquiring. The underlying representations, which do not appear in actual

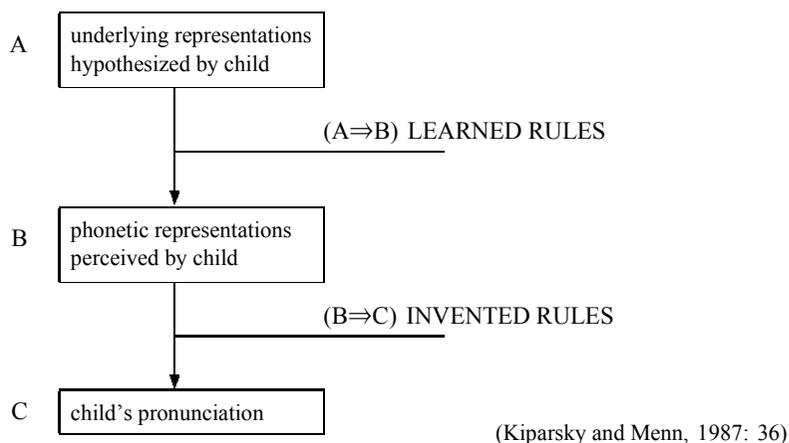


Figure 1. Kiparsky and Menn's framework

speech, are considered to form the knowledge of the phonological relationships between different forms of words. Stage B consists of the children's perceptions of the phonetic representations of the adult language that they are acquiring. The 'problem-solving' activity is conducted between this stage and the next stage, and the children's intended pronunciations are finally produced at Stage C. Their productions could be different in certain ways from the physical output, as when a purely physical limitation prevents them from producing segments that the children believe they are distinguishing.

According to Kiparsky and Menn (1987), in the early stage of language acquisition, Stages A and B coincide, but Stages B and C are completely distinct. As children master more of the phonetic repertoires of the adult language after the initial period of rule invention, Stage C approaches Stage B, and the system of rules shrinks. The process normally terminates when Stage C becomes identical with Stage B. On the other hand, as children continue to discover the phonological relationships of the adult language, Stage A becomes increasingly different from Stage B.

When Kiparsky and Menn's framework is applied to Second Language Acquisition (SLA), some crucial differences between FLA and SLA should be noted. Firstly, children start acquiring their L1s from scratch, whereas L2 learners have already acquired their L1 sound systems. Secondly, children normally have enormous opportunities for receiving input and feedback, but these opportunities are relatively limited in the case of SLA. Finally, not all L2 learners succeed in acquiring an L2, although all children basically do in FLA. Taking these differences into account, we will examine the acoustic data of the acquisition process of English vowels by Japanese learners, and then propose the acquisition model for L2 pronunciation.

3. The Summary of the Experiments on the Acquisition Process of the English Vowels

Isono (2000, 2003) investigated the acquisition process of the English vowels [ʌ, æ, ɪ, i:, ɛ] produced by 51 Japanese subjects consisting of three groups (the US, the PS, and the AS groups) and the native control group (the NS group) made up of 8 native speakers who had been teaching English in Essex, East Anglia, England. The characteristics of the three Japanese groups are summarised in Table 1 (See details in Isono, 2003).

Table 1. Summary of the US, the PS, and the AS groups

	US group	PS group	AS group
The number of the subjects	17 subjects (13 females : 4 males)	17 subjects (8 females : 9 males)	17 subjects (12 females : 5 males)
Status	Undergraduates	Postgraduates	Overseas students
Experience of studying abroad	Less than 1 month (2 of the subjects)	Less than 1 year (5 of the subjects)	More than 3 years (All of the subjects)
Major	Literature	Linguistics or Literature	Linguistics
Mean age	22	27	30

These Japanese groups were selected to reflect three learning stages of English, and as a result of the reading-aloud test, it was clarified that they were indeed at different levels in their pronunciation proficiency, which means, the most advanced Japanese subjects in order were the AS group, then the PS group, and the US group¹⁾.

The subjects were asked to produce five English words: *but*; *bat*; *bit*; *beat*; and *bet*, which contained the vowels [ʌ, æ, ɪ, i:, ɛ], in the carrier sentence ‘I say _____ again’, and their speech productions were recorded using a digital recording machine. Each target word was randomly shown at regular intervals to the subjects in all the groups three times. Therefore, a total of 885 different vowel productions (3 repetitions × 5 vowels × 59 subjects [{3 Japanese groups × 17 subjects} + 8 English subjects]) were obtained. Each of the 885 productions was characterised by the features of the frequencies of the first two formants (F1 and F2) and the vowel duration time by using the SUGI Speech Analyzer (produced by ANIMO Ltd.). The relationships between the frequencies of F1 and F2 and vowel sound quality are summarised as follows in Table 2.

Table 2. Relationships between frequencies and the characteristics of vowels

	F1	F2
High frequency	Low vowel	Front vowel
Low frequency	High vowel	Back vowel

Although various findings were observed in the experiment, the issue which is relevant to the current topic is that only the English [æ] was approximated to the norm as the Japanese subjects became more advanced learners. In other words, concerning the other English vowels [ʌ], [ɪ], and [ɛ]²⁾, approximations to the norms were not statistically observed regardless of the degree of the Japanese subjects' learning levels (See details in Isono, 2003).

Isono (2000, 2003) attributed the reason why different degrees of progress of approximation were found in the four vowels [æ] and [ʌ, ɪ, ɛ] to the view that the amount of L2 exposure made it possible for the advanced learners to produce a 'new' vowel (the English [æ] in the case of the study) in a native-like way, but it did not affect the productions of 'similar' vowels, such as [ʌ, ɪ, ɛ], following Bohn and Flege (1992) who investigated the English vowels [æ, ɛ, ɪ, i:] produced by two groups of adult German learners differing in the period of exposure to English. This view is based on Flege's (1987) assumption that a similar vowel which has a close counterpart in L1 is often regarded as being equivalent to the L1 vowel, but a new L2 vowel which does not have a counterpart in L1 can evade the process of equivalence classification if a sufficient input is provided to L2 learners. We previously outlined the phenomenon in L2 pronunciation that there are some L1 features more likely to be transferred than others, and noted this is because L2 learners themselves 'select' what to transfer and what not to transfer. This study should be respected in that it showed how the phonetic distance between the L2 sound and the corresponding L1 sound affects the selection process.

4. The Characteristics in the Acquisition Process of the English [æ]

In this paper, in addition to the above findings, we turn our attention to the accuracy of the English [æ] produced by the Japanese subjects and investigate what differences are found between the groups. Based on the results, we will find some characteristics in the acquisition process of the English [æ] which are particular to Japanese learners, and will examine one of the tendencies characterising the acquisition continuum of interlanguage phonology.

As previously mentioned, the English [ʌ] and [æ] produced by the Japanese subjects did not show the same rate of improvement, despite the fact that both of the English [æ] and the English [ʌ] are typically replaced by the Japanese [a]. In most cases, the English [æ] was more improved by the Japanese subjects as they became more advanced learners, but the English [ʌ] was not. In this respect, we could say that the first step for Japanese learners in acquiring the English [ʌ] and [æ] is to become aware of the difference between the Japanese [a] and the English [æ], and to make a distinction between the English [ʌ] and [æ] in their English productions by improving their English [æ]s, even

though their English [ʌ]s are still replaced by their Japanese [a]s. Therefore, the first analysis (Analysis 1) deals with how many subjects in each Japanese group were able to distinguish between the English [ʌ] and [æ] in their English productions by approximating their English [æ]s to the native-like English [æ], even though their productions may not be accurate enough.

Table 3 provides the details of Analysis 1, which examined how many subjects in each Japanese group met the above requirement³⁾.

Table 3. Details of the subjects who were able to distinguish the English [æ] from [ʌ]

	Number of subjects	Minimum difference	Maximum difference	Mean group difference
US group	2 / 17	338 Hz	441 Hz	389 Hz
PS group	9 / 17	212 Hz	703 Hz	435 Hz
AS group	11 / 17	239 Hz	663 Hz	403 Hz

Note. The minimum and maximum differences in this table are the values of the subject's mean differences of the three repetitions.

The Japanese subjects who were regarded as being able to distinguish the English [æ] from [ʌ] in their English productions are: 2 subjects from the US group; 9 subjects from the PS group; 11 subjects from the AS group, as Table 3 indicates⁴⁾. Since there were only 2 subjects in the US group, they were included in the PS group for the aim of the analysis as their personal scores in the test of reading-aloud did not reach the average score of the AS group and they did not have any experience studying abroad. This group is labelled the US/PS group in the analysis.

In the case of examining the English [æ]s produced by Japanese learners, the relationships between their intended English [æ, ʌ] productions and their Japanese [a] productions are mainly focused, because the former are typically replaced by the latter. However, in the case of other language learners, such as Portuguese or German learners, their English [æ]s are usually replaced by their [e]s in L1. Hence, some researchers have been interested in comparing these learners' intended English [æ] productions with their intended English [ɛ] productions (e.g., Major, 1987b; Bohn & Flege, 1992). In fact, the distribution of the English [æ] itself is closer to that of the English [ɛ] rather than that of the English [ʌ] (See details in Isono, 2003). Kent and Read (1992) also note the similarity of the formant patterns between the English [æ] and [ɛ] from the viewpoint of acoustic analysis. Therefore, as far as the acquisition process for the English [æ] by Japanese learners is concerned, the next step for the Japanese learners who can stop substituting the Japanese [a] for the English [æ] will be to control their articulations of the English [æ]s so that their intended English [æ]s are not misunderstood as the English [ɛ]s by native speakers.

Analysis 2, which is the main analysis in this paper, investigates how accurately the Japanese subjects who were not eliminated in Analysis 1 produced their English [æ]s, mainly by comparing their English [æ] productions with their English [ɛ] productions. This analysis focused on these 30 subjects divided into 3 groups (11 subjects from the US/PS group, 11 subjects from the AS group, and 8 subjects from the NS group), and the 180 vowel productions (30 subjects × 2 vowels × 3 repetitions) elicited from them were characterised by the frequencies of F1 and F2, and then the mean frequency values were calculated for productions of each subject.

Firstly, we focus on the characteristics in F2 which correspond to the frontness-backness dimension. Based on the results reported by Isono (2003) who investigated the acoustic differences between the English tested vowels and the corresponding Japanese vowels, the mean F2 frequency of the Japanese [a] was 1664 Hz and that of the English [æ] was 1828 Hz. Hence, when Japanese learners try to produce the English [æ] in a native-like fashion, they face the task of moving the whole raised shape of the tongue body to a more front position than the position needed for the Japanese [a]. Taking this fact into account, we see Table 4 showing the details of the group differences.

Table 4. Details of the group differences in F2 of the English [æ] and [ɛ]

	US/PS group		AS group		NS group	
	Mean (Hz)	SD	Mean (Hz)	SD	Mean (Hz)	SD
English [æ]	2177.4	244.4	2017.8	160.1	1828.5	53.8
English [ɛ]	2326.9	218.9	2287.4	108.4	2099.0	136.7

As the table presents, the Japanese subjects were successful in moving the whole raised shape of the tongue body to a more front position, but they exaggerated it too much, as long as the criterion is set on the British accent⁵⁾. Especially, the US/PS group's English [æ] has much higher frequencies than the NS group's English [æ], and it is even higher than the NS group's English [ɛ]. According to one-way ANOVA, the group difference for F2 of the English [æ] was significant, $F(2, 27) = 8.7, p < .01$, and Bonferroni's Post Hoc Tests reported that the US/PS group's English [æ]s had significantly higher frequencies than the NS group's English [æ]s at the .05 level.

The error bar on the next page shows each group's mean value for the difference between the English [ɛ] and the English [æ] ([ɛ] value minus [æ] value) and its distribution range (95% Confidence Interval for the Mean) in the frequencies of F2.

As the value for the NS group shows, the difference in the F2 frequency between the English [ɛ] and [æ] is approximately 250 Hz. This means that the F2 frequency of the English [ɛ] is about 250 Hz higher than that of the English [æ] and that the English [ɛ] is a more front vowel than the English [æ].

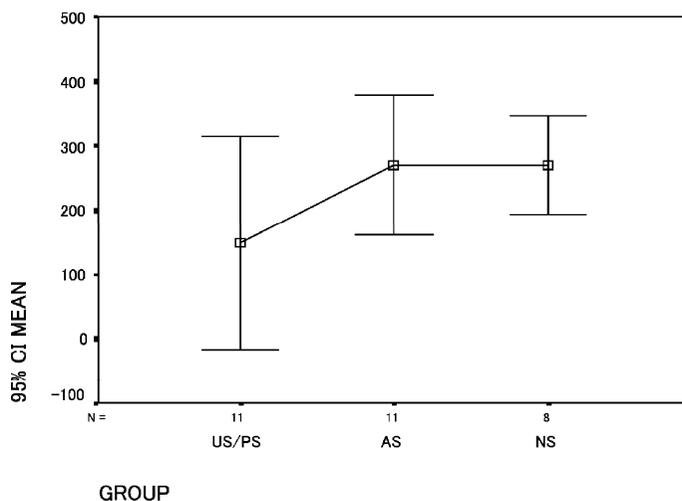


Figure 2. Differences between the English [æ] and [ɛ] in F2

Although the distribution range of the values of the AS group is slightly wider than the NS group, no noticeable difference is found in the mean differences between the two groups. However, when we pay attention to the US/PS group's value, we will notice that the mean difference is much smaller and the distribution range is much wider than the other two groups, although the differences for the mean values were not significant according to one-way ANOVA, $F(2, 27) = 1.4, p > .05$. It should be especially noted that the distribution range for the US/PS group even includes a minus value which means that some of the subjects in the US/PS group produced their English [æ]s more as front vowels than their English [ɛ]s. The question is why the subjects in the US/PS group exaggerated the 'frontness' of the English [æ], and it is a question to be considered later.

We now turn our attention to the F1 frequency which represents the vowel height dimension. The details for the acoustic data are summarised in Table 5. We can see that the F1 frequency for the US/PS group's English [æ] is much lower than that of the other two groups, and it means the US/PS group's English [æ] is closer to the English [ɛ] than the other two groups. One-way ANOVA revealed that there was a significant group difference for the F1 in the English [æ], $F(2, 27) = 4.7, p < .05$, and, according to Bonferroni's Post Hoc Tests, the US/PS group's F1 was significantly lower than the other two groups, at the .05 level.

Table 5. Details of the group differences in F1 of the English [æ] and [ɛ]

	US/PS group		AS group		NS group	
	Mean (Hz)	SD	Mean (Hz)	SD	Mean (Hz)	SD
English [æ]	829.2	146.8	924.3	110.9	989.7	43.6
English [ɛ]	655.9	82.2	647.4	88.3	695.0	57.1

The error bar presented below describes each group’s mean value for the difference in F1 frequencies between the English [æ] and [ɛ] and its distribution range. For the frequencies in F1, the value of the frequency of the English [æ] is normally higher than that of the English [ɛ], so the differences were calculated by subtracting the latter from the former.

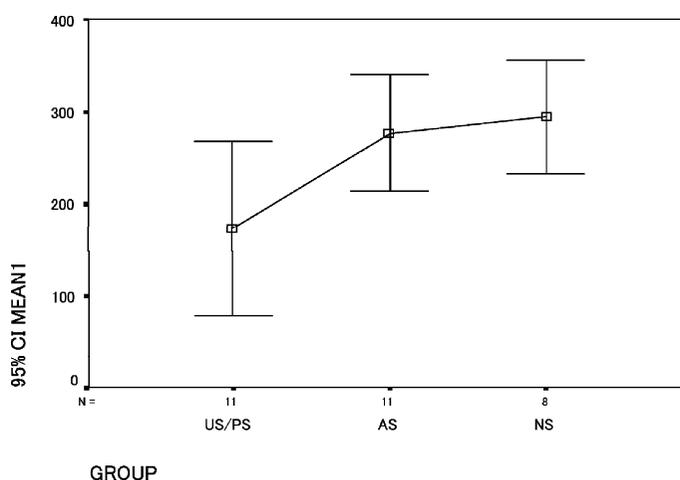


Figure 3. Differences between the English [æ] and [ɛ] in F1

As the figure shows, a similar tendency as was seen in the case of F2 is found. In short, a noticeable difference is not observed in either the mean difference or the distribution range between the NS and the AS groups, but the US/PS group’s mean difference value is much smaller and its distribution range is much wider than the other two groups. According to one-way ANOVA, a significant group difference for the mean difference was obtained, $F(2, 27) = 3.6, p < .05$, but Bonferroni’s Post Hoc Tests did not reveal a significant group difference at the .05 level. This is because the difference between the US/PS and the NS group was just not significant ($p = .07$).

We are now able to speculate about the acquisition process of the English [æ] for Japanese learners, if we regard the AS group as learners in the final stage of learning, the US/PS group as learners in the

intermediate stage, and the US group as learners in the beginning stage. In the early stage of learning, Japanese learners tend not to be able to distinguish the difference between the English [æ] and the Japanese [a], and directly substitute the Japanese [a] for the English [æ]. This was supported by the phenomenon that most of the subjects in the US group used the Japanese [a] as a replacement for their English [æ]s. As Japanese learners become more advanced learners, they stop substituting the Japanese [a] for the English [æ] as they become aware of the difference and try to produce their English [æ]s in a more native-like fashion. However, some of the Japanese learners' articulations are not accurate enough, and they produce the English [æ]s quite similar to their English [ɛ]s. This phenomenon was observed in both the vowel height and frontness-backness dimensions of the US/PS group's productions. This inaccurate pronunciation of the English [æ] will diminish when Japanese learners have more exposure to real spoken English as most of the subjects in the AS group did.

The reason why many subjects of the US/PS group, and actually some of the AS group, pronounced their English [æ]s similar to their English [ɛ]s can be attributed to the fact that Japanese learners are often instructed to produce an intermediate tone between the Japanese [a] and [e] when pronouncing the English [æ]. For example, Shimaoka (1994) has utilised *Katakana* for transcribing English segmental phonemes, and uses the symbol 「エア」 (= [ea]) as the representation of the English [æ]. The sizes of the letters correspond to the strength of the stress that learners are to use. In this method of transcribing, it is quite rare to find that a single English phoneme is transcribed by using two Japanese vowels, but this is due to the fact that Japanese has no counterpart for the English [æ]. Andoh (1984) presents the formula of 「(ア + エ) ÷ 2」 (= (a + e) ÷ 2), as a way for pronouncing the English [æ]. In Japan, this formula is a quite popular method for teaching the pronunciation of the English [æ], and it can be found not only in Andoh's book but also in some other textbooks (e.g., Ishiguro et al, 1992).

Figure 4 is an example of the US/PS group's English [æ] which might have been produced by following the above formulas. As the figure shows, the English [æ] completely diphthongises. The beginning part of the vowel is characterised by the sound which has approximately 2500 Hz in F2 frequency and 700 Hz in F1 frequency. This combination of F2 and F1 is quite similar to that of the Japanese [e], or even higher than that. However, after passing the midpoint of the vowel, the frequency of F2 decreases to around 2000 Hz, but that of F1 increases to 1000 Hz, which means the combination of F2 and F1 becomes closer to that of the Japanese [a].

The aim of these scholars' attempts is clearly to increase Japanese learners' awareness of the difference between the Japanese [a] and the English [æ] and to encourage them to stop substituting

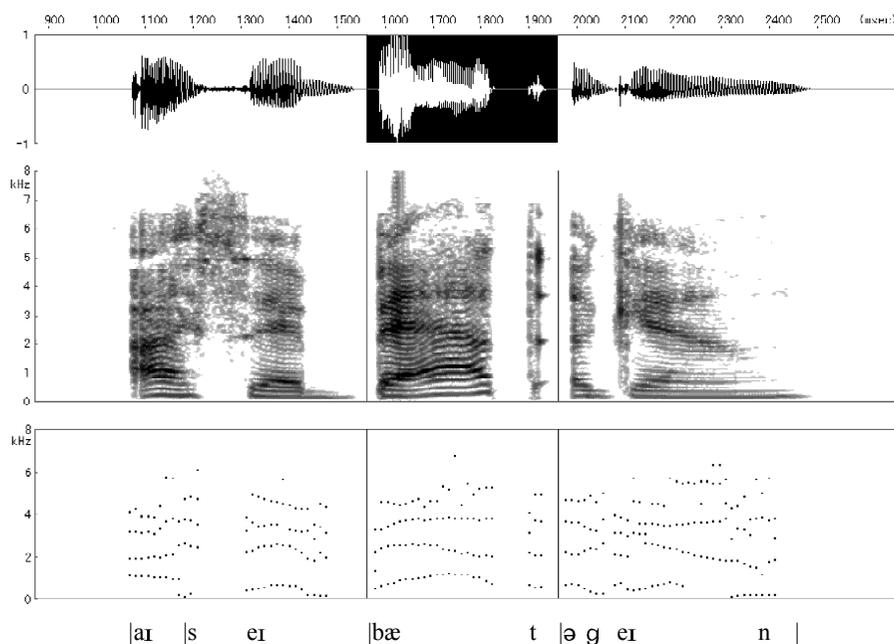


Figure 4. Example of formant patterns of the US/PS group' English [æ]

the Japanese [a] for their English [æ]s by attaching one more vowel to the Japanese [a]. However, when this method of instruction is used, it is quite natural that the attention of Japanese learners is directed at the existence of the Japanese [e] in these symbols and formulas. The reason is that mixing the Japanese [e] with the Japanese [a] is a crucial factor in allowing them to make a distinction between their intended native-like English [æ]s and the typically accented Japanese learners' English [æ]s which are directly replaced by the Japanese [a]. Therefore, by placing too much emphasis on the Japanese [e], Japanese learners stop the negative transfer of substituting the Japanese [a] for the English [æ]; however, their intended English [æ]s become quite similar to their Japanese [e]s and also to their English [ɛ]s. This would account for the phenomenon which was observed in the English [æ] productions of the US/PS group in this analysis.

5. The Acquisition Model of L2 Pronunciation

Taking the above findings and discussion into account, the framework shown in Figure 5 is proposed for the model of the acquisition process of L2 pronunciation.

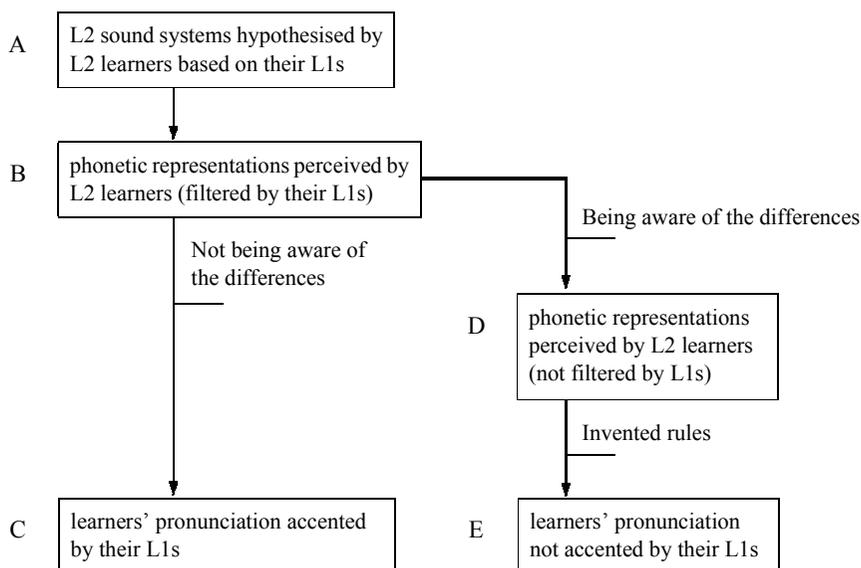


Figure 5. A framework of SLA in pronunciation

During Stage A, L2 learners only rely on the knowledge they have already acquired, so they substitute L1 sounds for L2 sounds. In Stage B, L2 learners receive input and feedback from formal instruction or by exposing themselves to the target language, but they still cannot escape from the L1 sound systems, because they hear the L2 sounds filtered through the L1 sound systems. L2 learners are not aware of the differences between L1 and L2 sound systems at these stages, as we can see from the fact that most of the subjects in the US group could not distinguish the difference the Japanese vowels and the corresponding English vowels, and substituted the former for the latter in their English productions.

However, as the result of enough input and feedback, some learners may become aware of ‘some’ of the differences between the two languages, and they come to perceive input they receive unfiltered by their L1s. This is Stage D. Concerning ‘some’, as previously noted, Bohn and Flege (1992) and Isono (2000, 2003) reported that it becomes easier for L2 learners to establish new phonetic categories as the phonetic distances between L2 and the corresponding L1 sounds are larger.

After passing Stage D, L2 learners will try to develop their L2 pronunciations so as to be more native-like, and this effort makes them invent their own rules and carries them forward onto the final stage, Stage E. As the English [æ]s produced by the US/PS group suggests, Stage D and E may be different especially at the early stage of acquisition because of the tongue’s muscular limitation or

wrong rules invented by L2 learners. However, as the acquisition stage proceeds, Stage E is assumed to approach to Stage D as we can see from the English [æ]s produced by the AS group, and the process terminates when they are identical. This process characterises the continuum of interlanguage phonology.

In contrast, L2 learners who do not become aware of the differences between L1 and L2 sound systems during Stage B will continue to substitute L1 sounds for L2 sounds, and will not proceed to the above route to Stages D and E. Their L2 pronunciation ends in Stage C, and is fossilised until they become aware of the differences.

This paper proposes the acquisition model of L2 pronunciation based on the experimental data of the English vowels produced by Japanese learners. This model may be useful for both of L2 learners and teachers. For L2 learners, it will provide an answer to the question about what kind of deviations L2 learners have to correct at the beginning stage. For teachers, it clarifies why L2 learners produce strange sounds named *interlanguage sounds* at the stage between Stage D and E, which belong to neither the L1 sound systems nor the L2 sound systems. Recognising this model, teachers will understand that the existence of interlanguage sounds in L2 learners' productions is not a bad sign, but it is a good sign, because these sounds become potential sources of native-like pronunciation.

Notes

1) The test material consisted of two texts taken from *MILESTONE: English Course I* (published by *Keirinkan*). It is a reading textbook for first-year high school students. The first text was directly reprinted from the textbook, but some words in the second text were changed by the researcher in order to meet the requirement that all English phonemes appeared at least twice, except for the English [ʒ]. The distribution of the English [ʒ] is quite limited, so it was very difficult to include the English words which include it in the two texts. The first text consisted of five sentences, and the second text was comprised of four sentences (See Appendix).

When the speech productions were edited by the researcher, only 3 native speakers' speech productions were randomly selected as the representatives of the NS group, and the other 5 native speakers' productions were removed from this reading-aloud test, in order to save time scoring the task. In addition, the first text was divided into two parts in order to control the length of the texts. Finally, the speech samples were blocked so that the same part for all subjects was on the same block. Naturally enough, they were randomised in each block with the requirement that the subjects in the same group did not occur back to back.

Each group of the 162 different stimuli (3 paragraphs × 54 subjects [51 Japanese and 3 English subjects]) was separately presented through a headphone, to each of 8 English native listeners who had teaching experience (M = 18 years). The order of the presentation of the blocks was B1, B2, and B3 for

all listeners. Their task was to determine the degree of the speaker's foreign accent on a 9 point scale from 'heavy foreign accent' (1 point) to 'no foreign accent' (9 points). They were encouraged to use the full range of the scale. They were told that they would be listening to the readings of Japanese and English speakers, but they were not told either the number of groups or that each speaker would be heard more than once. The timing of scoring was left to the listeners whether they would score after or during listening to each speech sample.

The foreign accent score for each subject was calculated as the sum of 24 scores (8 listeners \times 3 paragraphs). The foreign accent score for each group was calculated as the mean of subjects' scores in each group. The table below shows the results of the foreign accent scores and SDs of the Japanese and English groups.

Table 6. Foreign accent scores and SDs of the Japanese and English groups

	US group	PS group	AS group	NS group
Mean	122.2	150.8	172.8	215.66
SD	6.8	19.1	6.6	0.57

(Minimum Mean = 24, Maximum Mean = 216)

One-way ANOVA revealed that the group difference was significant, $F(3, 50) = 78.56, p < .001$, and according to Bonferroni's Post Hoc Tests, there was a significant difference in all cases at the .05 level.

- 2) Statistical results suggest that there is no significant difference in the shapes of the tongue among the English [i:], the Japanese [i], and the Japanese [i:] (See details in Isono, 2003).
- 3) It was quite a difficult task to establish a criterion for deciding whether Japanese subjects were able to distinguish the difference between the English [ʌ] and [æ] in their English productions. Consequently, the decision relied on the data from the NS group which was obtained in Isono (2003). Before outlining the criterion, some points, which were clarified in Isono (2003), should be noted in order to make the reasons for the decision clear. At first, it should be mentioned that there was not a significant difference among the English [ʌ, æ] and the Japanese [a] in the vowel height dimension. Therefore, in this first analysis, our focus will be limited to the characteristics of the frontness-backness dimension.

Concerning the distributions of the frequencies of F2 which represents the frontness-backness dimension, the Japanese [a] is midway between the English [ʌ] and [æ]. The English [ʌ] has generally a lower frequency, but the English [æ] has a higher frequency, in comparison to the Japanese [a]. This means, the highest frequency in the range for the NS group's English [ʌ] and the lowest frequency in the range for the NS group's English [æ] are the requirements to be met by Japanese learners in order for them to produce each of the English vowels in a native-like fashion. According to the data from the NS group which was elicited in Isono (2003), the highest mean frequency for the English [ʌ] was 1572 Hz, and the lowest mean frequency for the English [æ] was 1776 Hz, so the difference between them was 204 Hz. The mean frequency of the Japanese [a] elicited from the Japanese subjects in Isono (2003) was 1664 Hz, so the difference between the English [ʌ] and the Japanese [a] (the mean frequency of the Japanese [a] minus the highest mean frequency of the English [ʌ]) was 92 Hz. The difference between the English [æ] and the Japanese [a] (the lowest mean frequency of the English [æ] minus the mean frequency of the Japanese [a]) was 112 Hz. This argument is summarised in Figure 6.

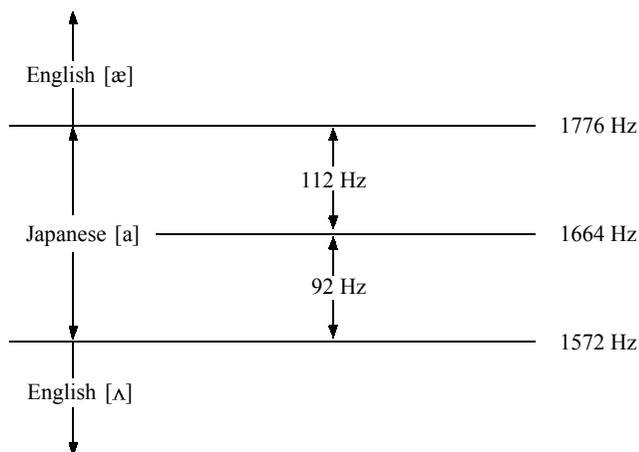


Figure 6. Differences in F2 among the English [æ] and [ʌ] and the Japanese [a]

Hence, when only the difference in the frequency is concerned, the value of 204 Hz is the criterion which decides whether a Japanese learner can distinguish between the English [ʌ] and [æ] in a native-like fashion. The values of 92 Hz and 112 Hz will be the criteria for judging whether the Japanese subjects are able to distinguish the Japanese [a] from the English [ʌ], and the English [æ] from the Japanese [a], respectively. Naturally enough, the frequency of the former vowel should be higher than that of the latter.

To sum up, for the purpose of making a distinction between the English [ʌ] and [æ], regardless of whether the Japanese subjects improve only their English [ʌ]s, only their English [æ]s, or both of them, their intended English [æ]s should be at least 92 Hz higher than their intended English [ʌ]s. Therefore, in this paper, only the Japanese subjects whose intended English [æ]s were at least 92 Hz higher than their intended English [ʌ]s in all three repetitions were regarded as being able to distinguish between the English [ʌ] and [æ] in their English productions.

As Table 3 shows, all the subjects whose intended English [æ]s were at least 92 Hz higher than their intended English [ʌ]s in all three repetitions necessarily produced their English [æ]s at least 212 Hz higher than their intended English [ʌ]s in the mean differences of the three repetitions. The fact that the number of subjects increases as they become more advanced is also consistent with the phenomena reported in Isono (2000). However, one thing which should be remembered is that these calculations are not related to the issue of whether they could produce the English [æ] or [ʌ] accurately, but are only based on the mean differences between the frequencies of their intended English [ʌ]s and [æ]s.

- 4) It was clarified that the rest who were excluded from the analysis substituted their Japanese [a] for their intended English [æ, ʌ] productions.
- 5) Even when we set the criterion on the American accent, their English [æ] still had much higher frequencies than the norm of American English, and the difference in the frequencies between their English [æ] and [ɛ] was quite small.

Appendix

Paragraph 1

Martin Luther King Jr. was a son of a black minister. His mother was a teacher. Young Martin spent a quiet childhood in Atlanta, Georgia. After high school, he went to college and studied to be a minister, like his father. In those days, who could guess he was going to be such a great leader of the black people?

Paragraph 2

Mary stayed at the town in the autumn to do some work. When she rented a cabin near the church, she asked for a boy to come and chop wood for the fireplace. One late afternoon, she found a boy standing at the door. His name was Arthur.

References

- Andoh, K. 1984. *Enshuu Onseigaku (A Practical Course in English Phonetics)*. Tokyo: Seibidou.
- Bohn, O. and Flege, J. 1992. The Production of New and Similar Vowels by Adult German Learners of English. *Studies in Second Language Acquisition*, 14(2): 131-158.
- Flege, J. 1980. Phonetic Approximation in Second Language Acquisition. *Language Learning*, 30: 117-134.
- Flege, J. 1987. A Critical Period for Learning to Pronounce Foreign Language. *Applied Linguistics*, 8: 162-177.
- Ishiguro, A. et al. 1992. *Jitsusen Eigo Onseigaku (Practical Phonetics)*. Tokyo: Kinseidou.
- Isono, T. 2000. Nihon-jin Gakushusha no Eigo Boin Shutoku Jyunjyo ni Kansuru Kenkyu (= The Study of the Acquisition Order of English Vowels by Japanese Learners). *The Bulletin of the Federation of English Education Society in Chuby-area*, 29: 87-94.
- Isono, T. 2003. *Japanese Learners' Interlanguage Phonology: with special reference to English vowels and plosives*. Unpublished Ph.D. thesis. University of Essex.
- Kent, R. and Read, C. 1992. *The Acoustic Analysis of Speech*. California: Singular Publishing Group, Inc.
- Kiparsky, P. and Menn, L. 1987. On the Acquisition of Phonology. In Ioup, G. and Weinberger, S. (eds.), *Interlanguage Phonology: the acquisition of a second language sound system*. New York: Newbury House.
- Major, R. 1987a. A Model for Interlanguage Phonology. In Ioup, G. and Weinberger, S. (eds.), *Interlanguage Phonology: the acquisition of a second language sound system*. New York: Newbury House.
- Major, R. 1987b. Phonological Similarity, Markedness, and Rate of L2 Acquisition. *Studies of Second Language Acquisition*, 9: 63-82.
- Shimaoka, T. 1994. *Chuukangengo no Onseigaku (Phonetic Acquisition of English through Interlanguage)*. Tokyo: Shougakukan.