# Improving Adult Japanese' Ability to Perceive English /l/ and /r/ via Self-Study with a Forced-Choice-and-Feedback Computer Game

**Antonio F. Smith**

## Introduction

It is commonly known that Japanese learners of English as a foreign language have difficulty acquiring the ability to distinguish between /l/ and /r/ (e.g., Goto 1971, Sheldon & Strange 1982, Nishiyama 1996, Moriuchi 1996, Cairns 1998). It is also commonly known that the difficulty increases with age (e.g., Miura 1996). Nonetheless, most Japanese started learning English in junior high school or later and never managed to fully acquire the two morphemes. It would therefore be potentially beneficial to a great many people, if a means were devised for older children and adults to fully acquire the categories /l/ and /r/. Prior laboratory studies (Logan et al. 1991, Lively & Pisoni 1994) have confirmed that Japanese adults do respond positively to forced-choice identification and feedback procedures, but none have established the procedures' practical limits. That is, none have determined whether native-like ability to distinguish /l/ and /r/ in all environments can be achieved. This pilot study is the precursor to a larger study that will test the effectiveness of the procedures over an extended time in a large body of adult Japanese students of English and should reveal whether or not the procedure can be used by students to achieve native-like ability in distinguishing between /l/ and /r/. Moreover, if the procedure is successful with the adult subjects, it is almost certain to be successful with children, given an appropriately motivating presentation, so Japanese teachers

of English could probably use it with students of any age.

## 2. Previous Studies

### 2.1. Lively and Pisoni (1994)

Studies have shown that laboratory procedures can influence adult Japanese perception of /l/ and /r/ (Logan *et al.* 1991). The most exhaustive study to date is by Lively and Pisoni (1994). One of their significant results was that subjects' accuracy improved, on average, by 12%, the average pre-test score of 65% climbing to 77% in the post-test.

In training, 19 paid adult (age 18-34) monolingual subjects first saw a minimal pair on a CRT screen, heard one member of the pair and then had to choose which member they heard. Subjects performed this task over three weeks, for a total of 15 training sessions, each lasting 40 minutes. Subjects heard one talker per training session pronounce a set of 68 minimal pairs twice. The total number of talkers was five–each being heard a total of three times. Improvements in accuracy occurred from week one to two and from two to three, in all five phonetic environments, #_V, C_V, V_V, V_C, and V_#, (albeit to a lesser degree in all but C_V in the latter period). Statistically, however, from week two to three, only the environments #_V, C_V, V_V improved significantly.

Nonetheless, in analyzing these results, it should be noted that the environments V_C and V_# were by far the most accurately identified, the former about 92% and the latter about 98%, judging from the publication's graphs, on average, in the third week. This suggests that as environments approach 100% accuracy, there is probably a tendency for the learning curve to flatten out. As the game for the proposed future study does not keep track of performance on individual environments, this tendency will probably result in a slowed rate of increase in scores over time.

Also, although it is tempting to conclude that there is a tendency for rate of improvement to decline over time in general, as four of the five environments declined

in the week two-to-three period, the rate of improvement in the environment C_V did, in fact, increase. Therefore, it is difficult to speculate about the sort of learning curve to expect from #_V, C_V, V_V over an extended training time, and unfortunately, as indicated above, the game used in the future study will not provide any information about this other than whether or not the environments can or cannot eventually be acquired.

## 2.2. Tsushima and Hamada (2000)

Another more recent study, by Tsushima and Hamada (2000), using similar training procedures but for seven contrasts–only two of which involve /l/ and /r/–/l-r/ and /gl-gr/–at syllable, word and sentence levels, found that perceptual training did significantly improve perception ability but that amount of training did not make a significant difference in the degree of improvement in subjects' perception ability.

Fortunately, this result of Tsushima and Hamada is corroborated by neither the Lively and Pisoni study nor, given some qualification, the author's pilot study. The different results may be due to differences in the training procedure and due to Tsushima and Hamada's focusing on a wider variety of contrasts than just /l/ and /r/.

The above-mentioned qualifications are that, in fact, two of the pilot's three subjects achieved some improvement but then did not improve significantly during the remainder of the training period, but there are plausible explanations for this–advanced age for one subject and inconsistency in training for the other. One of the pilot's subjects did achieve a consistent rate of improvement throughout the multi-month training period.

Finally, it seems possible that the Tsushima and Hamada subjects, who were asked to do perception training on their own, could also have been influenced by some subjects doing their daily listening homework inconsistently; moreover, some subjects could even have done the training improperly. If so, this might have influenced the

result that extended training did not produce a significant improvement over more limited training.

The present author's proposed study too will rely on students working alone, with individual subjects logging their own results over time, so the author will not be able to guarantee the accuracy of recorded scores, dates and times. It is possible that some students will not train regularly or even log false training sessions. Nonetheless, at least several students should train consistently, and most should log training accurately. Even if only ten out of seventy-five subjects train consistently, the future study should be a success, because those ten should reveal whether or not native-like proficiency is, in fact, achievable.

## 3. The pilot study

This pilot study has two aims :

1. To check whether or not subjects can in fact use the author's computer program to improve their ability to distinguish between /l/ & /r/
2. To provide data and constructive criticism with which the author can make an improved and larger study.

### 3.1. Method

#### 3.1.1. Subjects

The subjects were three native speakers of Japanese. All were volunteers and none were actively studying English at the time of the study. We will call the subjects H, N, and T. H is a seventy-year-old male with very little English. N is a forty-year-old female and an intermediate speaker of English. T, thirty-seven, is a highly advanced female speaker of English.

#### 3.1.2. Procedure

Subjects were given a copy of a computer program and asked to use it as

regularly as possible, keeping a log of their results.

### 3.1.3. The computer program

The author paid a computer programmer to make the program. It is in C++ and it is called "Speech Teach". The author recorded onto the program himself, another male, and two women reading a list of 200 minimal pairs prepared by the author. When a word started with /l/ or /r/, such as "lay" the speakers said, "The word lay", so as to put the initial /l/ or /r/ in a more natural environment, avoiding strong word initial stress. Then the individual words were "cut" out and labeled as individual voice files which the program could manipulate. From the subject's point of view, the program works as follows.

1. In a dialogue box, the subject sees the sentence, *Click "Play Word" and listen for a word to be pronounced, or choose "Quit"* and two virtual buttons. On one is written *"Play word"*; on the other is written *"Quit"*.
2. The subject clicks the button labeled *"Play Word"*.
3. The subject hears a word, for example, "lay", and the dialogue box changes.
4. In the new box is written, *"Highlight your choice:"* and below that is a minimal pair, here, "lay" and "ray" and two buttons. On one button is written *"Repeat word"* on the other, *"Submit choice"*.
5. The subject can highlight one member of the pair immediately and hit *"Submit choice"* or the subject can first hear the word repeated however many times s/he wishes by pushing *"Repeat word"* and then submit a choice.
6. The subject sees in a new dialogue box whether the choice was right or wrong. If correct, the subject sees, "That's correct!" the round number and two buttons, "OK", and "Show score"; if the subject hits OK he moves on to a new word (see 8. below). If incorrect, the subject sees, *"Sorry, the correct choice was 'lay'. Click 'OK' to continue, but listen carefully as the*

*word is repronounced...".* In this case, the round number and the two buttons, "OK", and "Show score" also appear.

7. The subject hears the word, here "lay", pronounced again.
8. The subject sees the original dialogue box again. And the whole process starts over with a new minimal pair, one member of which is pronounced. The pair and the pronounced member are both chosen randomly by the program.
9. After the tenth word, the subject sees a new dialogue box in which is written, *"That is correct!!/incorrect. This concludes Round 1. You have correctly chosen 8 out of 10. That's 80% accuracy."* Of course, the percentages vary. Here, the subject can click on *"OK"* and start a new round or first hit *"Show score"* to see what the total score is for all of the rounds of the current training session.

## 3.2. Results

The subjects kept track of the date, number of words heard and their score on Microsoft Excel. However, a score from any single training session might reflect high or low proportions of easy or difficult phonetic environments or more or less comprehensible speakers. To improve the chances of balanced stimuli, and therefore a more accurate result, scores from individual training sessions were grouped in fives (or as close to five as possible without grouping scores for sessions separated by a break of several days or more) and averaged. The average of the first five scores is compared to the average of the last. And the five-day averages are examined over time to see if the learning curve climbs steadily, flattens out or does something else.

At the end of the pilot study, subjects were asked to rank four phonetic environments by difficulty, according to notes about incorrect words as well as simple recollection. Subjects were not asked about [l-r]/V_C, as in "balk" vs."bark"

as few words of this type are included in the training set. All of the subjects agreed that the /l/ or /r/ in initial or final position was much easier than in medial position. N and T thought final position was easiest and H thought initial was easiest. All agreed that [l-r]/#C_V was more difficult than [l-r]/V_V. Apart from H's opinion about initial position, the subjects' opinions were in keeping with Lively and Pisoni's data (1994).

The two subjects who had rarely heard any of the speakers in the game use English agreed about the relative ease of comprehending the four different speakers.[2] The easiest to understand was a thirty-year-old female English teacher. The second easiest was a fifty-five-year-old woman. The third easiest was a fifty-seven-year-old male English teacher and the most difficult to understand was a thirty-one-year-old male English teacher, the author. All of the speakers spoke very clearly, except the author, who tried to alter his natural speech style as little as possible. That the subjects believed they more often chose correctly when the two female speakers pronounced the words might indicate that higher pitch contributes to comprehensibility.[2] Below are the significant numbers for each subject, followed by analysis.

### A-1 H's results

From 8/22/02 to 11/16/02, H used the program 60 times. His longest break between sessions was one three-day break; two-day and one-day breaks were fairly common. Each time he listened to 100 different words, and sessions lasted 10-15 minutes. The average of his first four scores was 48% correct. The average of the first five was 49%. The average of his last five was 55%. However, if we ignore the last three scores, which are low (perhaps because they come after a two-day break, or because of fatigue, the average is 56.8%. If we ignore the last four scores, the final five-day average is 57%. Using the 49% and 55%, his improvement is 6%. Using the 48% and the 57%, his improvement is 9%. Perhaps the most accurate figure is somewhere between the two, for example, 7.5%. If so, over 60 sessions, his rate of improvement is 0.125% per session, or about 79 minutes using the

program per 1 percentage rise in accuracy.

Assuming his final level is 56%, at the same rate of improvement, he would need 352 more sessions to reach 100% accuracy. Of course it is possible that his average rate of improvement might increase or decrease over time.

## A-2 Analysis

Two significant factors to keep in mind when considering H's modest rate of improvement relative to the other two subjects (as you will see below), despite his being the most fastidious, are his age, seventy, and motivation to learn /l/ and /r/, which is almost nil, apart from his kind desire to do the author a favor. H is not studying English now and does not plan to in the future. Nonetheless, he did improve his five-session average by 6%, or 8% if we ignore the last four sessions, which at first glance seems promising. Looking at the scores over time, however, groups of five averaged together, it will be seen that H's accuracy improved for a while and then leveled off. This result shows that there may indeed be an age limit for acquiring new phonetic categories. The overall trend for H's scores to remain roughly flat between 50% and 60%, after an initial modest increase, can be seen by the graph below.

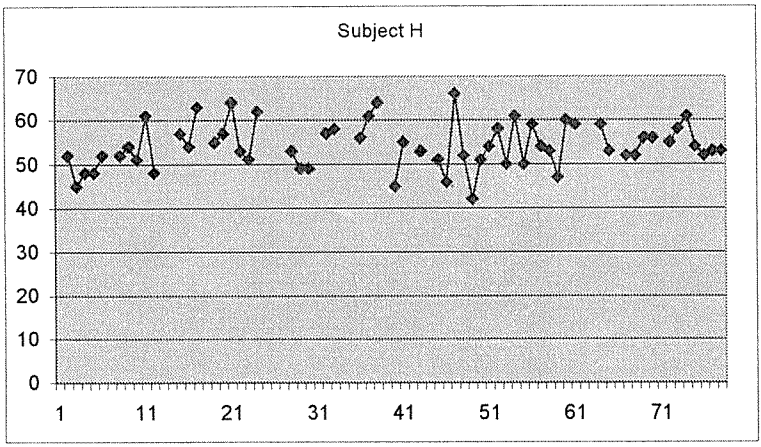

**FIG 1. Subject H's results**

| Section | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Average | 49 | 53 | 57 | 55 | 54 | 56 | 51 | 55 | 53 | 56 | 54 | 56 |

**Table 1. Subject H's results in terms of five-session averages over the training period**

### B-1 N's results

From 8/22/02 to 10/28/02, N used the program 48 times. Her longest breaks were one eight-day break and two three-day breaks. She almost always listened to 60 words at each session, and sessions lasted about 10 minutes. The average of her first five scores was 56%. The average of her last five is 73.6%. That is a 17.6% improvement, which equates to 0.366% per session or a 1-percentage point improvement for every 27.27 minutes using the program. At the same rate of improvement, it would take N 72 more sessions to reach 100% accuracy.

### B-2 Analysis

N's higher English proficiency, and younger age, 40, probably helped her achieve a higher rate of improvement than H. Probably, she also had higher motivation than H, as she teaches Japanese to foreign students and must sometimes communicate in English. She might perceive a practical advantage to being able to distinguish between /l/ and /r/. Her good consistency, except for a completely understandable eight-day trip to Thailand, undoubtedly contributed to her impressive results. When she did take a break, however, we can see obvious drops in her scores, and it takes a few sessions to get back to the same average performance she had before a break. Further, in contrast to H, her five-day-average scores rise steadily over the entire training period. They show no evidence of flattening out. This result shows that

it may very well be possible for adults of her age to acquire new phonetic categories.

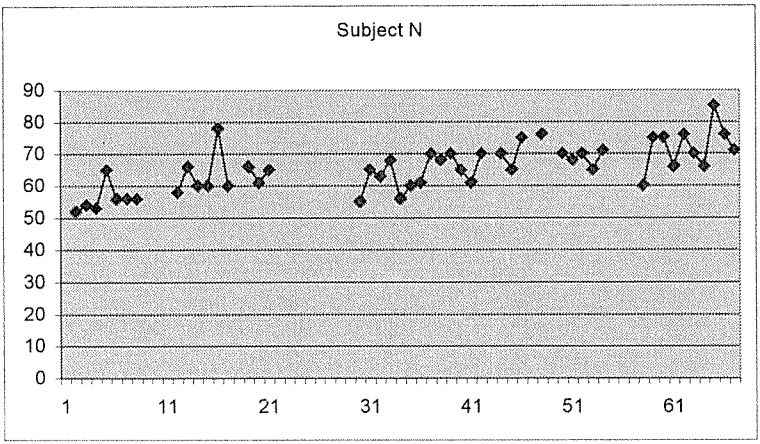


**FIG 2. Subject N's results**

| Section | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Average | 56 | 59 | 65 | 63 | 63 | 67 | 71 | 68 | 71 |

**Table 2 Subject N' s five-day averages over the training period**

The five-day averages stop three days short of the end of training as N had 48 training days (see FIG 2). However, the last three days of training were among her best scores; recall that her average for the final five days (not shown in the table) is 73.6%.

### C-1 T's results

T used the program 24 times during the training period, from 8/24/02 to 11/12/02, but she was very inconsistent, as can be seen in FIG 3, so it does not make sense to take five-day averages, as was done with the other subjects. Instead, we have

to look at groupings of data in an ad hoc way. Therefore, please refer to FIG 3 for the following discussion.

T's longest breaks are 16 days, 10 days, 8 days, 6 days and 5 days.[3] She almost always listened to 80 different words in a session. Sessions lasted 12-15 minutes. The average of her first 3 scores was 73%, the first four 76% (perhaps some learning had taken place) and the first five (after an 8-day break) 75.6%. Probably her starting level is somewhere between 73% and 76%. Her final 5-day average is 79%, for a total gain of just between 6% and 3%, but her last score is mediocre, comes after a 16-day break with just one intervening session, and may be related to her using headphones for the first time, which took some getting-used to. Therefore, the last session should be excluded. In addition, the aforementioned intervening session resulted in a low score (76%) undoubtedly because of the ten-day break before it, so it should be excluded as well.

Then, if we look at the last four legitimate scores, the average is 81.5%. But, according to T's journal, on the last of those four sessions she was extremely tired but tried to do the program anyway. Therefore, if we exclude the score for that session (10/22) and take a three-day average ending 10/20, the result is 83.6% (the score of 10/17 is also excluded due to a note about extreme tiredness in T's journal, thus the three-day rather than four-day average).

This three-day average score of 83.66% is preceded by a twenty-one day stretch in which T used the program only three times. Prior to that long period of infrequent practice, T had her most consistent period, 11 sessions in fifteen days. This is the period I will focus on in analyzing T's results in the following section. The average of the last two scores in that period is 84%.

I think the scores of 83.6% and 84% represent T's peak performance as a result of using the program. The average of the last six scores, excluding sessions after long breaks and sessions that suffered due to extreme tiredness, is also 83.66%.

The score of 84% was reached after just 15 sessions. If her true initial level

is somewhere between 73% and 76%, for example 74.5%, then her increase is 9.5% in 15 sessions, or 0.63% improvement per session, averaged over the 15 sessions that took place in 29 days. Remarkable! Assuming 13.5 minutes per session, T. achieved an average increase of 1 percentage point in accuracy for every 21.3 minutes using the program during the 15 session training period.

If she had maintained the consistency of her first 15 sessions over the remaining sessions and that consistency had produced the same average rate of increase of 0.63% per session, she would have reached 100% accuracy in 40 days. Of course, it is very unlikely that she would maintain that rate of improvement. Probably, the learning curve would start to flatten out after she mastered the phonetic environments that came most easily to her.

## C-2 Analysis

T's progress from 8/24 to 11/17 suffers because of inconsistency. Even her impressive two-day average score after her first fifteen sessions might have been higher had the first fifteen sessions not included an eight-day break.

Compared to her scores on sessions just *before* long breaks, her scores on sessions just after the breaks dropped considerably. Truly rapid improvement only occurred during the aforementioned 11-session stretch over fifteen days. Tomoko begins with 4 sessions in six days, takes an eight-day break and then begins her 11-session stretch. The average of the first two sessions in that 11-session stretch is 76% correct out of 160 words (the same as her initial four-day average, before the eight-day break), and the average of the last two sessions is 84% correct out of 160 words, as mentioned above.

I think that once T gets fully warmed up again, she should be close to the 84% mark. If so, and if she could be persuaded to begin using the program consistently, T would make an interesting candidate for further study. She would begin at about the accuracy level that the best Lively and Pisoni subjects had when they finished training,

so it would be interesting to see how her learning curve would behave above 84%. It would reveal to what extent at least one Japanese adult can acquire the /l/ and /r/ distinction by using the high variability forced choice method.

On the other hand, T hears a great deal of English from her American husband, friends and professors.[4] This exposure to English may have given her a latent advantage in benefiting from the forced choice program.

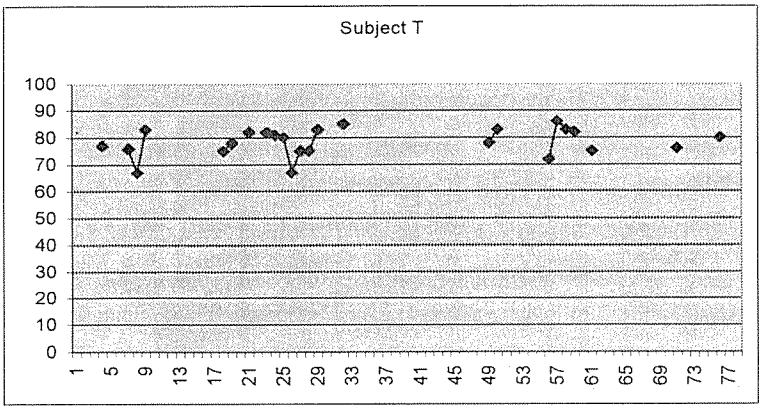


**FIG 3. Subject T's results**

### 3.3. Constructive criticism of program from subjects

There are two primary criticisms, one involving feedback, and the other motivation. The criticism regarding feedback is that after being informed of whether their submitted choice was right or wrong, subjects wanted to be able to again have access to the *"Repeat word"* button, or even be able to contrast the members of the pair by clicking on them in turn.

Regarding motivation, subjects wanted more in the way of rewards for improved scores. For example, subjects suggested more praise using a greater variety

of words and expressions for correct choices. They also wanted graphics, such as increasing numbers of stars, fireworks or dancing people set to increments of 60%, 70% 80% correct etc. They also suggested music of increasing length and/or different type to go with increasing scores.

When asked, subjects replied that they did not want the program to keep track of their performance in the different phonetic environments and allow them to work on problematic environments. They said the initial low scores would be discouraging.

## 4. The improved larger study

The larger study is an attempt to transform theoretical linguistics into applied linguistics. Its purpose is to test how well essentially the same procedure used by Lively and Pisoni works with adult students of English studying mainly on their own.

However, its results should prove interesting to theoretical linguists. Lively and Pisoni studied 19 students for three weeks and got an average of improvement in accuracy of 12%, an increase from 65% to 77%. From this result, they conclude, "Monolingual native speakers of Japanese can learn to identify English /r/ and /l/". I think they are overstating their case. It is entirely possible that the learning curves of subjects will flatten out over time–the harder phonetic environments proving insurmountable obstacles–with learners never attaining100% accuracy.

Fortunately, given the number of subjects, their motivation, and the length of the training period, the proposed future study has a very good chance of revealing whether or not non-native learners can in fact achieve native accuracy using the procedure. Moreover, the pilot study's training method, which will also be used for the larger study, appears more efficient than the method used by Lively and Pisoni. The total number of minutes of training in the Lively and Pisoni study is 600, which equates to an overall average rate of improvement of 1% for every 50 minutes of training. That is better than H's rate of 79 minutes per 1% increase; but the age of

their subjects is 18 to 34 to his 70. The other two pilot study subjects, age 37 and 40, though older than those of the Lively Pisoni study, improved at roughly twice the speed of the Lively and Pisoni subjects. This faster rate of learning increases students' chances of reaching their maximum possible accuracy level within the time of the study.

## 4.1. Method

### 4.1.1. Subjects

The subjects of the proposed larger study would be the three first-year English major classes at Osaka University of Foreign Studies–approximately 75 eighteen-to-nineteen-year-old advanced and, presumably, highly motivated learners of English. The proportion of female to male students is about three or four to one.

Their youth gives them a decided advantage over all three pilot study subjects, as does their motivation. On average, however, they will have much less exposure to English than the pilot subject T, but much more exposure than H and N. Their rate of improvement per session could easily be somewhere between N's and T's, at least up to the level of 84% accuracy or so. Thereafter, there is no precedent for rate of improvement.

### 4.1.2. Revisions to the game

At this point, probably the only revision to the game before the larger study would be some means for a word to be pronounced repeatedly after a choice has been submitted and the result (correct/incorrect) reported.

### 4.1.3. Procedure

A revised version of the game software would be given to each student and all students would be told to use it for at least 12 minutes four times a week outside of class. We would use an amplified version of the game in class for about twelve

minutes once a week, and students would be given special sheets for recording their choices, which would be collected after the class session. Students would also be given a log sheets to keep track of the dates of sessions and scores outside of class, and the logs would be collected once a month.

Students would be shown the results of the pilot study in graph form and told to plot their progress on a similar graph that would be given to them. This way, they could get a mental picture of their average rate of increase over time, and they would not be discouraged by short-term up and down fluctuations.

### 4.1.4. Training Period

To estimate the amount of training needed for students to have a chance of reaching 100% accuracy, one must first estimate their starting percentage of accuracy and their rate of improvement. First of all, the pre-test scores of Lively and Pisoni's "monolingual" subjects, that revealed an initial accuracy of 65%, seem strangely high when compared to H's initial score of 49%. Although the pilot subject H knows many English words, he is indeed close to being monolingual. Even the pilot subject N, an intermediate English speaker, had a starting score of only 56%. Only the highly advanced English speaker, T., had a higher initial score, 76%. Is it possible that an A.B.D. Ph.D. student in linguistics at a highly ranked American University has a score that is only 11% better than true Japanese monolinguals? Given the discrepancy between the initial scores in the Lively and Pisoni study and those of the pilot study, it seems reasonable to assume that the Lively and Pisoni subjects were not really monolinguals.

The discrepancies might be partially explained by differences in stimuli in the two studies. For example, it could be the case that the "talkers" on the Lively and Pisoni pre-test pronounced words much more carefully than the speakers on the pilot's game. Moreover, it is certainly true that the listening conditions of the game are not nearly as good as those of the laboratory used in the Lively Pisoni study. Still,

it seems unlikely that such differences could fully account for the observed discrepancies. Probably, the Lively and Pisoni subjects modestly underestimated their English level when reporting it to the researchers.

As it is better to err on the part of caution, it would be safer to assume that the subjects of the proposed study will have an initial average score somewhere between N's 56% and T's 76%. A plausible estimate might, in fact, be about 65%.

As the average age of the subjects will be 18, it is very possible that their rate of learning will be even better than that of N and T, but conservatively assuming a rate of improvement of a little less than that of N and T, for example, a 1% gain for every 30 minutes on the program, students would need 17.5 hours to reach 100% accuracy. However, as the learning curve might well flatten out to some degree, it would be best to add to the training time.

Unfortunately, the amount of time students can be expected to use the program is limited. I assume a reasonable request is that students use the program five times per week, four times outside of class and one time in class, for at least twelve minutes per time, which yields a total of one hour per week and a total of 15 hours in the first semester. At this level, students should not feel too inconvenienced.

Students will be advised, however, to use the program more often and longer for better results, and some students will probably do so. In the event that the training appears to be working but that more time is necessary to achieve native-like proficiency, students can be asked to use the program during the semester break and even during the second semester.

## 5. Conclusion

The results of the pilot study promise even more interesting results for the proposed future study. The training method has been shown to be effective. Subject N's accuracy increased steadily for the entire training period, and subject T's accuracy increased steadily whenever she used the program regularly. Only subject H's scores

showed no appreciable increase in the latter part of the training period, probably due to his age, 70, and lack of practical motivation for learning the categories. The subjects for the future study, however, are even better candidates for improvement than N and T in terms of age and motivation. They are adults, but just barely, and their motivation should be high as they are English majors and know their chances of getting a good job related to their major will be influenced by their English ability. Finally, the amount of time available for training is large, up to a full academic year. It is, therefore, very likely that at least some students will use the program consistently and for enough time either to succeed at reaching native-like proficiency or to prove the limitations of the training method and/or adult learners.

## References

Cairns, R. 1998. Taking the long road ; or was it the wrong load? : Japanese learners and the l/r problem. *Kobegakuin daigaku jimbungakubu kiyo* 17 : 123-32.

Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9 : 317-23.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura Y., and Yamada, T. 1994. Training Japanese listeners identify English /r/ and /l/. III. Long term retention of new phonetic categories. *Journal of acoustic society of America* 96 (4) : 2076-87.

Logan, J. S., Lively, S. E., and Pisoni, D. B. 1991. Training Japanese listeners to identify English /r/ and /l/ : A first report. *Journal of acoustic society of America* 89 : 874-86.

Nishiyama, T. 1996. Nihonjin eigo hatsuon no ayamari bunseki. *Kyoikugaku kenkyu kiyo* (chugoku shikoku kyoiku gakkai) 42(2) : 170-3.

Miura, I. 1996. Discrimination of segmental and suprasegmental phones by Japanese students leaning English from an early age. *International review of applied linguistics in language teaching* 34 : 135-43.

Moriuchi, N. 1996. Nihonjin eigo gakushusha no onsei no shutoku jokyo nitsuite. *Kyoikugaku kenkyu kiyo* (chugoku shikoku kyoiku gakkai) 42(2) : 187-92.

Sheldon, A. and Strange, W. 1982. The acquisition of /r/ and /l/ by Japanese learners of English : Evidence that speech production can precede speech perception. *Applied psycholinguistics*. 3(3) : 243-61.

Tsushima, T. and Hamada, M. 2000. Effects of intensive perception and production training on improvements of nonnative speech perception ability in a university English course. *Collected articles on the English language*, Part 6 (TEFL) 33 : 982-93.

Acknowledgements: Much thanks to Joel Jones for volunteering to program at highly discounted rates out of his interest in the project. Much thanks to William Nelson of Osaka Gaikokugo Daigaku, Cindy Daugherty of Kyūshū Kōgyō Daigaku and Linda Wallace of the University of California at Santa Cruz for volunteering to have their voices recorded. Much thanks toUeda Isao of Osaka Gaikokugo Daigaku for helping with the making of the individual voice files. Thanks to the subjects who shall remain anonymous. And thanks to Elisa Jones for assisting with the list of minimal pairs.

1 The third subject is the author's wife and has regular exposure to some of the speakers on the program but not others, so I will not consider her results, which appear skewed because of this.
2 The most comprehensible in the Lively and Pisoni study (1994) was also female.
3 Her circumstances during the time of the study made regularity difficult.
4 She is a Ph.D. candidate at an American university.