A Preliminary Study on Stimulus Order Effects in Discrimination of the English/b/-/v/Contrast during Categorial AX Discrimination Training.

Teruaki TSUSHIMA

1. Introduction

Previous research has demonstrated that, though perception of speech segments is generally categorical, it is still susceptible to auditory, phonetic, phonemic or more global contexts in which it occurs (Repp, 1987; Repp & Liberman, 1987). As one of such contextual effects on speech perception, it has been shown that discriminability of a speech contrast significantly changes when the order of stimuli is reversed in some discrimination procedures (Bohn & Polka, 2001; Cowan & Morse, 1986; Crowder, 1982; Francis & Ciocca, 2003; Healy & Repp, 1982; Polka & Bohn, 2003; Repp & Crowder, 1990; Repp, Healy, & Crowder, 1979). The effects are generally called "stimulus order effects" or "presentation order effects".

Early studies investigated the stimulus order effects in discrimination of vowels using an AX discrimination procedure (Cowan & Morse, 1986; Crowder, 1982; Healy & Repp, 1982; Repp & Crowder, 1990; Repp, *et al*., 1979). Cowan & Morse (1986), for example, tested discrimination of the English vowel contrast i/-/ɪ/ by adult native speakers of English, using a synthetic vowel continuum from /i/ to /ɪ/. The AX discrimination procedure with the ISI (inter-stimulus-interval) of 0, 250, 500, 1000 and 2000 ms was employed. It was found that discrimination accuracy was significantly higher when the more peripheral vowel served as the first stimulus than when the order was reversed. It was also found that the magnitude of the order effects was greater as the ISI increased. Based on these results, they proposed a hypothesis that the memory code of the first stimulus shifts toward a neutralized vowel during a silence period, making discrimination relatively difficult.

The following study by Shigeno (1992) tested discrimination of the English /i/-/ɪ/ contrast by adult native speakers of Japanese. The identification data showed that both the English /i/ and /ɪ/ were within a category boundary of the Japanese /i/, although the English /ɪ/ was shown to be a perceptually more deviant instance of the Japanese /i/ than the English /i/. The results showed that the order effects were found at the /i/-end of the continuum, but not at the

/ɪ/-end. She suggested that the order effects were not obtained at the /ɪ/-end because the listeners were not able to utilize phonetically categorized memory codes, but turned to the auditory information in discriminating the stimuli. The results suggested that phonetic (or phonemic) categorization might play an important role in the order effects, and that the order effects may result from a language-specific (as opposed to general auditory) processing of speech stimuli.

This position gained support from a more recent study by Francis & Ciocca (2003). They reported that, in discrimination of Cantonese lexical tones, adult native speakers of Cantonese showed higher sensitivity when the first token in a pair had a lower F_0 (low-high order) than otherwise, but that English-speaking counterparts did not show such effects. Moreover, neither group showed the order effects in discrimination of non-speech analogues that had the same F_0 as the Cantonese speech stimuli. The study suggested that, in the AX discrimination procedure, listeners might utilize linguistically relevant categories whenever they were available, and that the interaction of the memory codes of the stimuli were conducive to the order effects.

The stimulus order effects gained a recent attention in the literature from a developmental perspective (Bohn & Polka, 2001; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Bohn, 2003). Polka & Bohn (2003) in reviewing the literature on the infant speech discrimination studies, found that discriminability of a speech contrast was generally lower when the vowel used as a background was more peripheral in a vowel space than when the order was reversed, regardless of the phonemic identify of the target vowels. They proposed a peripherality hypothesis that "the perceptual asymmetries point to a language-universal perceptual bias that infants bring to the task of vowel discrimination such that the relatively more peripheral vowel in a contrast serves as a reference vowel (p. 224)." They further hypothesized that these asymmetries might go through developmental changes through experience of listening to the native vowel categories such that the initial biases toward the peripheral vowels may be retained, enhanced, or reduced according to the structure of the native vowel categories. According to the hypothesis, the order effects observed across developmental stages of listeners might reflect a structure of perceptual categories in the listener's perceptual space.

The results and suggestions provided by the previous research led to a hypothesis that stimulus order effects would be obtained in discrimination of nonnative speech contrasts. In order to test this hypothesis, the author and his colleagues conducted a series of studies, focusing on discrimination of the English consonant contrasts by Japanese listeners (Tsushima, 2007a, 2007b, 2007c; Tsushima, Sasaki, Shiraki, & Yoshida, 2001; Tsushima, Shiraki, Yoshida,

& Sasaki, 2003, 2005; Tsushima, Yoshida, Shiraki, & Sasaki, 2003, 2005）. The first focus of the studies was discrimination of the American English /b/-/v/ contrast. It is known that Japanese listeners normally assimilate both sounds of the contrast into the Japanese phoneme /b/. The English /b/, however, is perceptually closer to the Japanese /b/ than the English /v/. Another contrast that has been studied is the American English contrast /l/-/r/. There is evidence that the English /l/ is perceptually closer than the English /r/ to the Japanese /r/ into which both sounds of the contrast are assimilated（Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004）. It was hypothesized that such an asymmetry in perceptual distance of nonnative speech sounds to their corresponding native sound would produce significant order effects in the AX discrimination procedure.

Supporting the hypothesis, the studies found that discriminability was significantly lower when the first stimulus was a nonnative sound that was perceptually closer to the corresponding native sound（*i.e.*, "/b/ to /v/" or "/l/ to /r/"）than otherwise（*i.e.*, "/v/ to /b/" or "/r/ to /l/"）in certain listening conditions under the AX discrimination procedure（Tsushima, 2007c; Tsushima, Shiraki, *et al.*, 2005; Tsushima, Yoshida, *et al.*, 2003, 2005）. Specifically, significant order effects were obtained under a condition where listeners were required to utilize some form of categorized information in the discrimination task. Tsushima *et al.*（2003）, for example, examined the order effects in discrimination of /b/-/v/, manipulating the ISI and stimulus variability[1)]. Significant order effects were obtained at the ISI of 1500 ms, but the effect was a lot weaker at the ISI of 300 ms（see Tsushima（2007c）for similar results on /l/-/r/）. In addition, significant order effects were observed in the discrimination task where an irrelevant vowel was inserted between the stimuli, as well as in the AX identification procedure（Tsushima, Shiraki, *et al.*, 2005; Tsushima, Yoshida, *et al.*, 2005）. These results suggested that the effects might be based on some kind of categorized information in memory that might be utilized in the discrimination task（*e.g.*, a phonemic or context label）.

Finally, the results found some evidence to suggest that the order effects might be due to an assimilation process（Tsushima, Shiraki, *et al.*, 2005; Tsushima, Yoshida, *et al.*, 2005）. For example, it was found that, among the incorrect responses when the order of presentation was "/b/ to /v/", the proportion of responses for "/b/ to /b/" was significantly greater than that of the other responses（*i.e.*, "/v/ to /b/", or "/v/ to /v/"）. It was hypothesized that, when a nonnative sound that was perceptually closer to the native sound（*i.e.*, /b/）was perceived and stored in the memory, it might perceptually assimilate the following perceptually less close sound（*i.e.*, /v/）at the time of discrimination.

The author and his colleagues also examined the order effects during discrimination train-

ing, rather than in an initial exposure to the stimuli (Tsushima, 2007b; Tsushima, Shiraki, *et al.*, 2003). In Tsushima, Shiraki, *et al.* (2003), for example, the initial test on discrimination of /b/-/v/ was followed by discrimination training with feedback. Six sessions of standard AX discrimination training with 9 blocks of 16 trials were held using feedback in terms of correctness of their responses. Following the training, a posttest was administered to check for the improvements of identification and discrimination ability. The results showed that only a non-significant tendency of the order effects was observed at the beginning of the training sessions.

Tsushima (2007b) specifically examined whether the order of stimulus presentation in the AX discrimination training had any significant effects on improvements of the ability to identify and discriminate the American English /l/-/r/. One group of participants (the /l/-first group: $N=16$) received the training with the stimulus order of "/l/ to /r/" or "/l/ to /l/" while the other group (the /r/-first group: $N=15$) were trained with the order of "/r/ to /l/" or "/r/ to /l/". Four sessions of categorical AX discrimination training were conducted with feedback in terms of correctness of their responses. In order to encourage listeners to categorize the stimuli across different phonetic contexts, lexical items were used for the speech stimuli. Fourteen lexical items, produced by four native speakers of English, were used (*e.g.*, *light-right*) for the speech stimuli. It was hypothesized that the assimilation process, which was induced by presentation of /l/ as the first stimulus, would have detrimental effects on improvements of discrimination and identification ability.

The results did not support the hypothesis, however, showing that the degree of improvement did not significantly differ between the /l/-first group and the /r/-first group. The /l/-first group even surpassed the /r/-first group in the identification test at the posttest. The results might well be due to the low level of discrimination performance. The English /l/-/r/ contrast is known to be one of the most difficult contrasts to perceive for Japanese listeners (Tsushima, 1999; Yamada, 1995). In addition, the use of lexical items required the listeners to categorize the stimuli across differing vowel contexts. The average percents correct for "different" stimuli (*i.e.*, "/l/ to /r/" or "/r/ to /l/") in both groups were not greater than 65% at the posttest. It might be the case that the difference in the stimulus order did not make significant contributions to discrimination learning when the participants performed only slightly above the chance level.

The present preliminary study was designed to extend the previous studies on the effects of stimulus order during AX discrimination training, with the following major modifications. First, it focused on discrimination of the English /b/-/v/ contrast, which is known to be easier to perceive by Japanese listeners than the English /l/-/r/ contrast (Brown, 1994). Second,

speech stimuli consisted of nonsense syllables（rather than lexical items）with the same vowel（*i.e.*, /ba/ and /va/）. Finally, the design of the training was modified such that participants were required to pass four stages that were incrementally more difficult. It was believed the changes in the design would encourage participants to maintain their attention to the discrimination tasks through the end of the training sessions.

The specific question asked in this study was, "Did the stimulus order significantly influence improvements of the ability to discriminate the English /b/-/v/ contrast by Japanese learners of English during categorial AX discrimination training?"

2. Method

2–1. Participants

Participants were 19 monolingual Japanese students at Ryutsu-Kagaku University in Kobe, Japan, who were at the basic or pre-intermediate level in terms of their English ability. None of them reported any history of hearing or neurological impairments. Participants were randomly divided into the /b/-first group（$N=10$）and the /v/-first group（$N=9$）, according to the stimulus order used in the discrimination training.

2–2. Speech Stimuli

Speech stimuli were identical to those used in the author's previous studies（Tsushima, *et al*., 2001; Tsushima, Shiraki, *et al*., 2003, 2005; Tsushima, Yoshida, *et al*., 2003, 2005）. They were natural tokens produced by two male native speakers of English. The speakers recorded multiple tokens of stimuli, using a frame sentence, "I will say X". Among them, two tokens for each category per speaker were selected such that non-critical acoustic cues to the contrast did not systematically differ between the two categories（*e.g.*, VOT, F_0 and duration）.

2–3. Procedure

2–3–1. Training Program

The training program consisted of a practice block and the following four consecutive stages of discrimination training（*i.e.*, First, Bronze, Silver and Gold Stage）. Each stage was composed of multiple blocks, each of which had 12 trials. Participants could pass a certain stage if the percent correct was greater than a predetermined criterion in two consecutive blocks. The criterion for First, Bronze, Silver, and Gold Stage was 60%, 70%, 80% and 90%, respectively.

The program started with general instruction with regard to how the participants should complete the training by passing all the four stages, as well as what they were asked to do in the discrimination task. Then, they were asked to launch the program, type in necessary information (*e.g.*, the participant's name) on the computer and initiate the training program. When they reached the criterion in one block at a certain stage, a message appeared on the monitor saying that they would go to the next stage if they would surpass the criterion in the following block. When they passed a certain stage, a message congratulated them for having done so, and instructed them to go to the next stage. The training was terminated 20 minutes after the initiation of the program.

2–3–2. Experimental Procedure

A categorial AX discrimination training procedure was used for discrimination training. In this procedure, participants were asked to indicate whether a pair of the stimuli belonged to the same or different category by pressing "s" for the "same" response, and "d" for the "different" response on the keyboard. In half of 12 trials in one block, two different tokens from the same category (*e.g.*, b1–b2) were paired (*i.e.*, "same" trials), while in the other half, those from different categories (*e.g.*, b1–v2) were paired (*i.e.*, "different" trials). In each group, the order of stimuli was "AA" (*i.e.*, "/b/ to /b/" or "/v/ to /v/") in 6 trials, and was "AX" (*i.e.*, "/b/ to /v/" or "/v/ to /b/") in the other 6 trials in one block. The order of presentation with respect to the stimulus order was randomized within a block. The ISI was set at 1200 ms in order to encourage the participants to utilize categorized information for a stimulus comparison, as opposed to lower-level information such as acoustic differences. After the second stimulus was presented, the participants had two seconds to respond. When the response was correct, feedback was immediately provided on the monitor for 500 ms in the form of a red circle. A red cross (*i.e.*, "x") was used as a negative feedback when the response was incorrect or no response was made within the time limit. The next trial began two seconds after the offset of the visual feedback.

The training program was created by the author using a stimulus presentation software (*Superlab: Cedrus Corporation*), which controlled the presentation of the audio and visual stimuli, and recorded the participants' responses and reaction times. On the day of training, each participant sat at a computer in a quiet room, wearing a high-quality headset which was hooked up to an amplifier. The audio stimuli were presented at a comfortable level of approximately 65 *dB*.

3. Results

3–1. The Number of Participants Who Passed the Gold Stage

The first analyses examined the differences between the /b/-first group and the /v/-first group in the number of participants who completed the training program by passing the Gold stage within the time limit. It was found that only one participant in each group failed to complete the training within the time limit. The participant in the /b/-first group had a great deal of difficulty in the First stage, while the one in the /v/-first group in the First and Bronze stage. Thus, the results did not show any significant differences between the two groups in terms of the number of participants who completed the discrimination training.

3–2. The Number of Trials Required to Pass Each Stage

The next analyses compared the number of trials required to pass each stage between the two groups. As is shown in Figure 1, the mean number of trials was much greater in the /b/-first group in the First stage, Silver stage, and Gold stage. Examination of individual data, however, showed that the difference in the mean number of trials was largely due to the fact that a few participants in the /b/-first group needed a disproportionately large number of trials to pass these stages. For example, there were three participants in the /b/-first group who needed 180, 132 and 132 trials, respectively, to pass the First stage, while the other participants had little difficulty passing it (*e.g.*, 24 trials)[2]. A statistical analysis using a non-parametric test (*i.e.*, Mann-Whitney test) showed that the difference in the number of trials was not significant in any of the stages ($p > .05$). In sum, no significant differences were found between the two groups with regard to the number of trials required to pass each stage. The only suggestion provided by the data was that discrimination might prove particularly difficult for a limited number of participants in the /b/-first condition, especially at the beginning of the training.

3–3. Percents Correct as a Function of Stage

Next, the percents correct in each stage were compared between the two groups. As is shown in Figure 2, participants in both groups performed well above the criterion in the First stage (*i.e.*, 60%), Bronze stage (*i.e.*, 70%), as well as in the Silver stage (*i.e.*, 80%), indicating that the discrimination tasks were already relatively easy in the First stage. A comparison of the

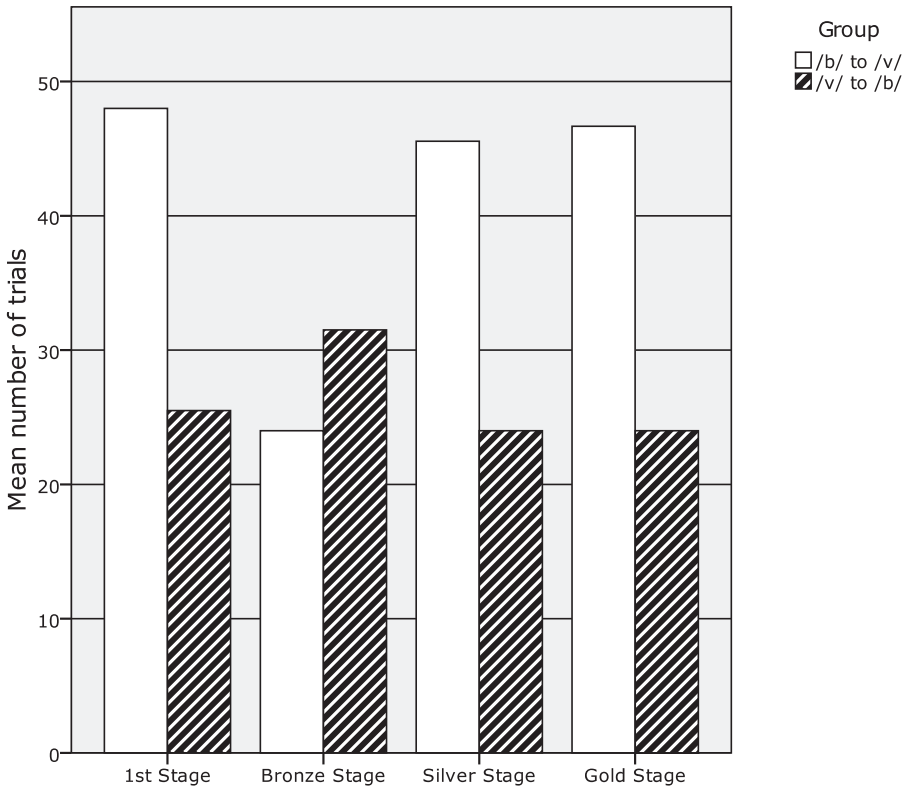

Figure 1 Mean number of trials as a function of group and stage.

two groups at each stage showed few notable differences between the /b/-first and /v/-first condition.

In summary, the present results did not support the hypothesis that presentation of a nonnative sound closer to the corresponding native sound as the first stimulus would have detrimental effects on nonnative discrimination learning. The following analyses compared the participants' performances in different stimulus order conditions within each group, as well as across the two groups.

3-4. Percents Correct as a Function of Stage and Stimulus Order

Figure 3 shows the mean percents correct as a function of stage and stimulus order in the /b/-first group. First, it is shown that the percent correct was substantially lower in the /b/ to /b/ than in /b/ to /v/ condition in the First stage, while the direction was reversed in the following stages. The percent correct in the /b/ to /b/ condition dramatically improved to over 90% in the Bronze stage, while that of the /b/ to /v/ condition gradually increased from the

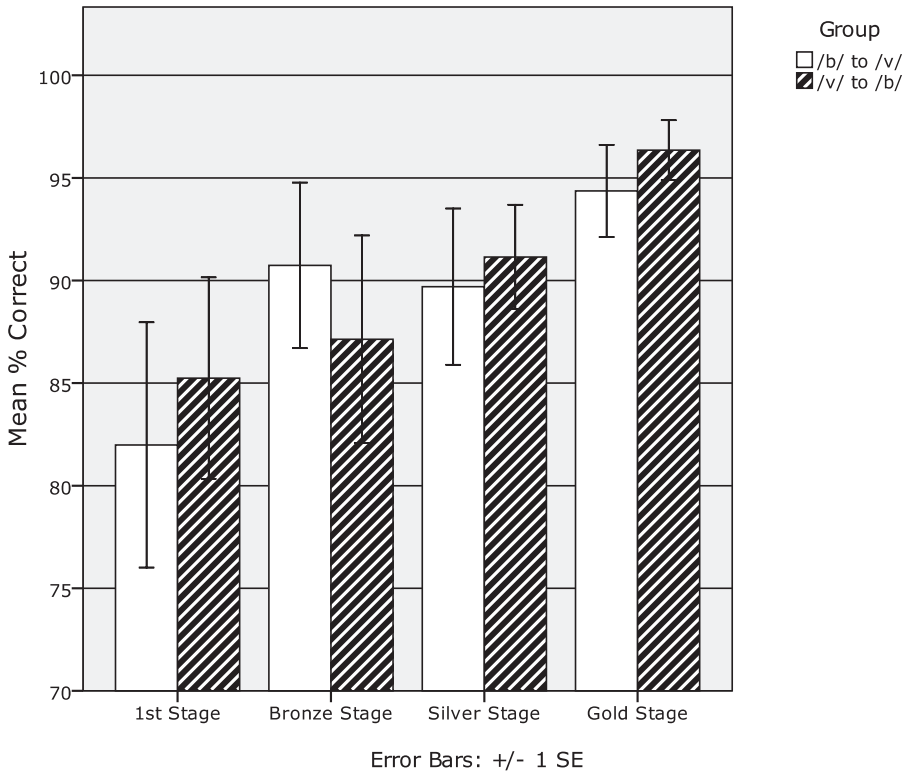

Figure 2 Mean percent correct as a function of group and stage.

First to the Gold stage. Examination of individual data showed that, in the First stage, four out of nine participants performed less well in the /b/ to /b/ than in the /b/ to /v/ condition. Especially, the participant who needed a great number of trials（*i.e.*, 132 trials）to pass the First stage described in 3-2. above, performed disproportionately worse in the /b/ to /b/ than in the /b/ to /v/condition（*i.e.*, 25.6% versus 83.3%）. A Wilcoxon singed-ranks test showed that the difference in the percent correct between the /b/ to /b/ and /b/ to /v/ condition in the First stage did not reach significance, $Z = 1.753$, $p = .08$.

Figure 4 shows the mean percents correct as a function of stage and stimulus order in the /v/-first group. As is shown in the figure, the percent correct was consistently higher in the /v/ to /v/ than the /v/ to /b/ condition across the stages. Examination of the individual data showed, however, that performance differences between both order conditions were not very substantial. The percent correct in the order of /v/ to /b/ was lower than that of /v/ to /v/ in only three and four participants（out of ten）in the First and Bronze stage, respectively. The difference in the mean percent correct between the two conditions was not statistically significant in any of the stages（$p > .05$）.

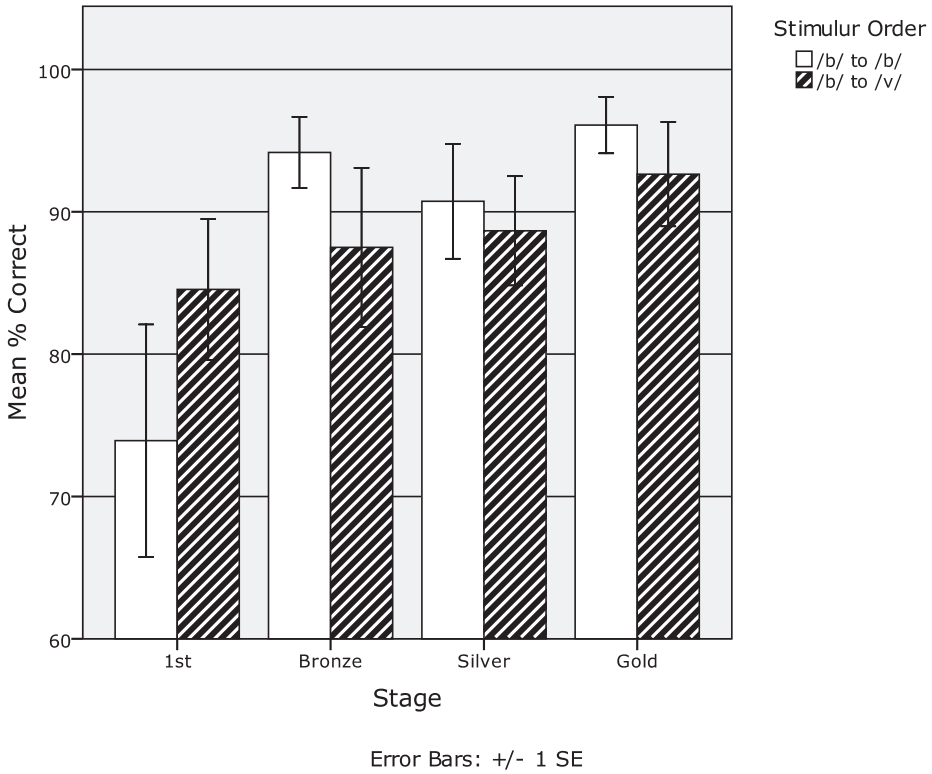

Figure 3 Mean percent correct as a function of group, stage and stimulus order in the /b/-first group.

Finally, the percent correct in the /b/ to /v/ and /v/ to /b/ condition was compared across the two groups. A comparison of Figure 3 and Figure 4 showed no indication that the percent correct in the /b/ to /v/ condition was significantly lower than that of the /v/ to /b/ condition across the stages. On the contrary, the percent correct was actually lower in the /v/ to /b/ condition in the first three stages. A comparison of the /b/ to /b/ with the /v/ to /v/ condition showed that the percent correct was substantially lower in the /b/ to /b/ than in the /v/ to /v/ condition in the First stage due to the particularly low percent correct in the /b/-first group. The difference, however, turned out not to be statistically significant in any of the four stages ($p > .05$).

In sum, the analyses of the percents correct as a function of stage and stimulus order did not find any significant differences in the level of performance between the /b/ to /b/ and /b/ to /v/condition in the /b/-first group, or between the /v/ to /b/ and the /v/ to /v/condition in the /v/-first group. There were no significant differences between the /b/ to /v/ and the /v/ to /b/ condition, and between the /b/ to /b/ and /v/ to /v/condition, either. The results revealed the unexpected tendency that the performance was worse in the order of /b/ to /b/

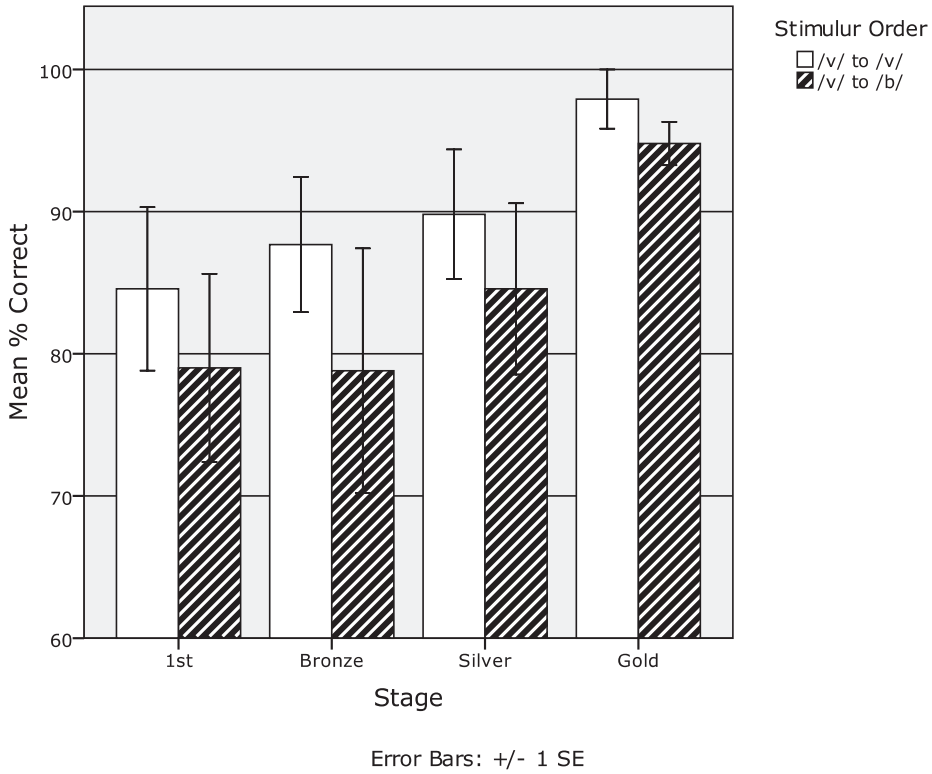

Figure 4 Mean percent correct as a function of group, stage and stimulus order in the /v/-first group.

than that of /b/ to /v/ in the First stage, suggesting that some mechanisms other than the perceptual assimilation of nonnative sounds to the native sound might operate under the present conditions.

3–5. Reaction Times as a Function of Stimulus Order

Figures 5 and 6 show the mean reaction times as a function of stage and stimulus order in the /b/-first and the /v/-first group, respectively. First, the mean reaction times substantially decreased across the stages in both groups, as expected. This indicated that the training facilitated the perception of speech stimuli as well as the decision processes required to perform the discrimination tasks. Second, the mean reaction time was substantially longer in the /v/-first group than in the /b/-first group in the First stage, but that the difference largely disappeared in the Gold stage. Finally, the mean reaction time was slightly longer in the /b/ to /b/ condition than in the /b/ to /v/ condition in the /b/-first group in the first three stages, while no definite pattern was observed in the /v/-first group. A repeated-measures analysis of variance

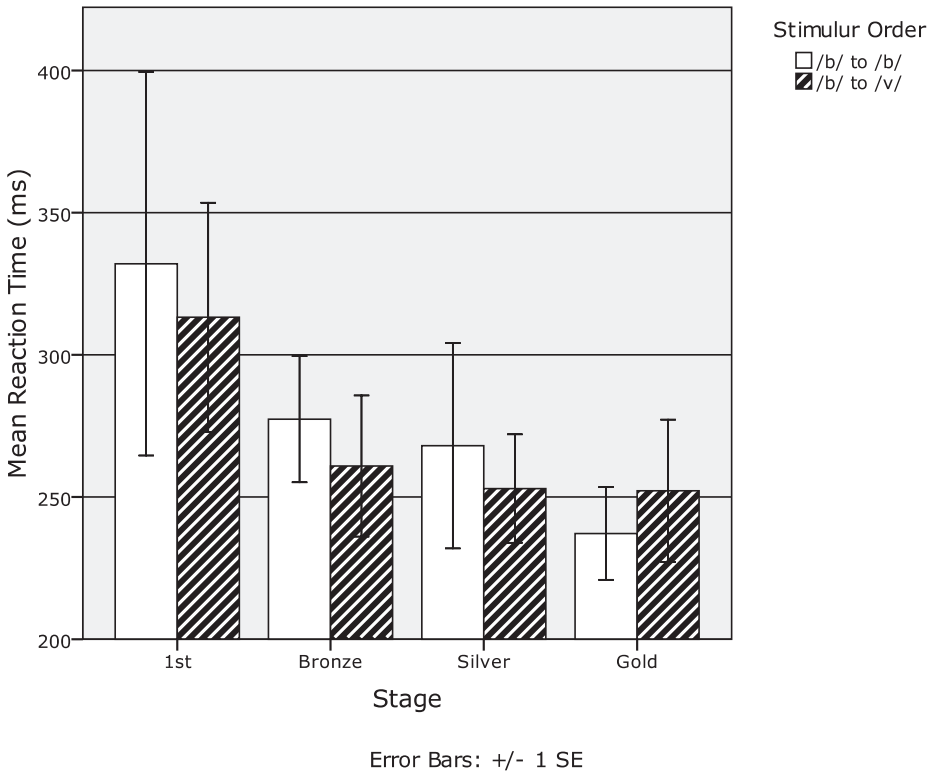

Figure 5 Mean percent correct as a function of group, stage and stimulus order in the /b/-first group.

(ANOVA) was conducted with the group (/b/-first, /v/-first) as the between-subject factor, and the stage (First stage, Gold stage), stimulus order (the same category (*i.e.*, /b/ to /b/ or /v/ to /v/) and the different categories (*i.e.*, /b/ to /v/ or /v/ to /b/)) as the within-subject factors. The main effect of stage was significant, $F(1, 15) = 9.39$, $p = .008$, indicating that the reaction times significantly decreased from the First stage to the Gold stage, averaged across the groups. Post-hoc tests with Bonferroni-adjusted pair-wise comparisons ($p < .05$) showed the following. First, the mean reaction times significantly decreased in the /v/-first group in both /b/ to /v/ and /v/ to /b/conditions ($p = .008$ and $p = .004$, respectively), while the difference was not significant in either condition in the /b/-first group, largely due to relatively large individual variability. Second, the difference in the reaction time between the /v/ to /b/ and /b/ to /v/ condition was marginally significant in the First stage ($p = .072$). Finally, none of the effects of stimulus order was significant within each group at any stage.

In summary, the results did not support the hypothesis that longer reaction time was required to perform the discrimination task in the /b/-first than in the /v/-first condition. On the contrary, the reaction times tended to be longer in the latter than in the former condition at

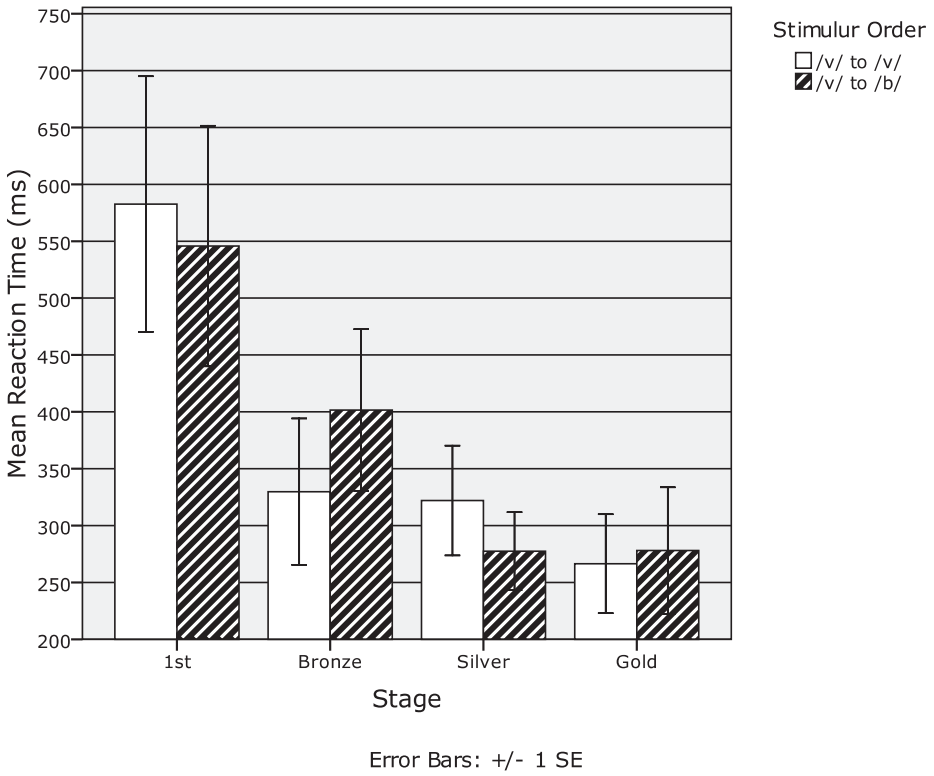

Figure 6. Mean reaction time as a function of stage and stimulus order in the /v/-first group.

the beginning of the training. Within the /b/-first group, there was no evidence that longer reaction time was required in the /b/ to /v/ than in the /b/ to /b/ condition.

4. Discussion

The present study attempted to test the hypothesis that the stimulus order in categorial AX discrimination training would have significant effects on improvements of the ability to discriminate the English /b/-/v/ contrast. It was hypothesized that the improvements would be negatively influenced under a condition where /b/ was presented as the first stimulus（*i.e.*, /b/-first condition）as compared with the condition where the stimulus order was reversed（*i.e.*, /v/-first condition）.

The present results clearly did not support the hypothesis, showing no significant differences in the degree of improvements in the discrimination ability between the /b/-first and /v/-first condition in terms of either the percent correct or reaction time. It was found, however, that a limited number of participants showed a great deal of difficulty in the /b/-first con-

dition, suggesting that the effects of the particular stimulus order condition might have some limited effects on discrimination learning. The results were compatible with Tsushima (2007c) which showed that discrimination learning was not negatively influenced when /l/ was used as the first stimulus in discrimination of the English /l/-/r/ contrast during categorial AX discrimination training.

Nor did the present results support the hypothesis that an assimilation process is involved in the order effects in discrimination of /b/ and /v/. A strong tendency was found that performance was worse in the /b/ to /b/ than in the /b/ to /v/condition at the First stage in the /b/-first condition. This result was surprising because the opposite direction of the effect was expected as the participants could have easily categorized both stimuli into /b/ and have correctly responded "the same" under this condition. First, it might be the case that the participants incorrectly categorized the stimuli based on some non-critical acoustic cues (*e.g.*, voice onset time) such that different tokens of /b/ were given different labels. Second, it is also possible that the participants had a strong response bias toward responding "different" when they were not sure of the phonemic identity of the stimuli. At any rate, the results indicated that the hypothesized assimilation process, if any, might not be induced under the present condition.

The results of the present study, combined with the findings of previous studies, support the conclusion that significant order effects hardly occur under a condition where participants are trained to discriminate the stimuli with feedback, while they can be obtained under a condition where participants are initially exposed to nonnative speech stimuli and tested without feedback (Tsushima, Shiraki, *et al.*, 2003, 2005; Tsushima, Yoshida, *et al.*, 2005). First, the absence of significant order effects during training might be due to some differences between the initial exposure and the discrimination training in the way nonnative speech sounds are perceived. In the initial exposure to the nonnative speech stimuli, listeners might be forced to perceptually assimilate nonnative sounds into perceptually similar native phonemes because of relatively high stimulus uncertainty. It might be under this condition that the assimilative process in the /b/-/v/ order is induced in the discrimination process. During the training, on the other hand, listeners are likely to attempt to find the critical acoustic cues by changing the native-language attentional patterns into those required to distinguish the nonnative contrast. Accordingly, the order effects may fade away at the beginning of the training sessions as they learn to correctly perceive the stimuli by attending to the critical cues. Neurophysiological research that compares brain activities in the initial exposure and during discrimination training might provide important information regarding this hypothesis (cf. Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005).

Second, another possible reason for the absence of significant order effects during training might be attributable to some changes in the discrimination strategies between the initial exposure and the training. In one discrimination strategy, for example, the participants may give a phonemic label to the first stimulus, and use it as an anchor when the second stimulus arrives. In the other strategy, they may give a context label（*e.g.*, "close to /b/"）to the first stimulus without specifying the phonemic identity, and then give phonemic labels to both stimuli when the second stimulus arrives. The hypothesized assimilation process might be induced in the former strategy as the first stimulus is phonemically categorized and stored, supposedly being assimilated to a perceptually similar native phoneme. It would be interesting, in future research, to examine how instructions to participants regarding the discrimination strategy might influence the direction and magnitude of the stimulus order effects.

Finally, the failure to obtain significant order effects, particularly in the present study, might be accounted for by the relatively low level of difficulty in the discrimination tasks. As shown in Figure 2 above, the mean percent correct in both order conditions surpassed 80 % even in the First stage. The low level of task demands was probably due to the following factors. First, the phonemic category of the first stimulus in the pair was fixed such that it was always /b/ in the /b/-first group or /v/ in the /v/-first group. Although not informed of this design, most of the participants might not have needed a long time to notice that the phonemic identify of the first stimulus was fixed, except for a few who had a great deal of difficulty under the "same" condition. Second, the target phoneme appeared in a nonsense syllable（rather than a lexical item）where the following vowel was always /a/. As was pointed out in Introduction, this measure was taken because the lexical discrimination training conducted in Tsushima（2007c）turned out to be too demanding for participants. It, in turn, appeared to make the tasks in the present experiments much less demanding.

Another issue that was addressed in the present study was whether participants' attention to the task could be kept during the discrimination training. The training adopted a design where participants were required to pass multiple stages with increasingly difficult criteria. According to the author's observation, most of the participants were able to keep their attention to the task until they passed the Gold stage. It was especially felt that the design helped them keep their attention to the task because of the motivation to pass each stage. This observation may obtain indirect evidence from the reaction times. As is shown in Figures 5 and 6, the reaction times continued to decrease until the Gold stage, suggesting that the participants attempted to perform more efficiently until the end of the training.

The results of this preliminary study suggested some important modifications to be made

in the following full-scale study. First, in order to make the level of task difficulty in discrimination learning higher, more stimulus variability should be introduced. In the present study, the stimuli consisted of four tokens for each category, which were produced by two different speakers. It would be necessary to increase the number of speakers such that listeners would discriminate the stimuli produced by different pairs of speakers across the stages. For example, they would discriminate the stimuli produced by a pair of Speaker A and B in one stage and those by a pair of speaker C and D in the other stage. Second, the following study should use a pretest-posttest design in which participants' identification abilities are tested before and after the training, using different sets of stimuli in the training and the pre/posttests. It is certainly important to examine whether learning that occurs during discrimination training is generalized to new stimuli and a new procedure. This would provide important data on whether discrimination training enables learners to develop perceptual categories in the long-term memory that can be utilized in identifying the nonnative speech sounds.

5. Conclusion

The present study attempted to examine whether the stimulus order of the English /b/ and /v/ in the AX discrimination training procedure had significant effects on discrimination learning by adult Japanese listeners. It was found that that the stimulus order conditions did not have significant effects on the overall improvements of discrimination ability. First, the number of participants who failed to complete the program within a time limit did not significantly differ between the two conditions. Second, the number of trials required to pass each stage did not significantly differ, although some tendency was observed that more trials were needed in the /b/-first than in the /v/-first condition. Finally, no significant differences were observed between the two conditions in the mean percent correct or the mean reaction time in any stage.

The present results, combined with those of the previous studies, point to a conclusion that significant order effects in discrimination of nonnative speech contrast are obtained under some limited listening conditions; 1) listeners are initially exposed to speech stimuli; 2) ISI is relatively long or very short (*e.g.*, 2000 ms, 100 ms); 3) stimulus variability is relatively high (*e.g.*, the multiple-talker condition). Under these conditions, relatively strong perceptual assimilation of nonnative speech sounds into perceptually similar native sounds is likely to occur. Future research may be directed toward specifying the mechanisms underlying the stimulus order effects, by further examining various aspects of experimental factors that determine the

magnitude and direction of the effects.

注

1）The ISI was varied in order to examine whether and how the order effects were affected by the kind of the memory available in the discrimination task（*e.g.*, auditory, phonetic and phonemic）. Stimulus variability was varied to examine whether the order effects were influenced by the degree of stimulus uncertainty in the task. In a single-talker condition, one stimulus for each category was used in one condition while two stimuli each from two different talkers were used for each category in a multiple-talker condition.

2）In the Silver stage, one participant needed 206 trials to pass the stage, while in the Gold stage, two participants needed 108 and 144 trials, respectively.

Acknowledgements

This research was generously supported by a research grant from Ryutsu-Kagaku University in the academic year of 2008.

References

Aoyama, K., Flege, J. E., Guion, S., Akahane-Yamada, R., & Yamada, T.（2004）. Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233–250.

Bohn, O. S., & Polka, L.（2001）. Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *Journal of the Acoustical Society of America*, 110（1）, 504–515.

Brown, C.（1994）. The role of the L1 grammar in the L2 acquisition of segmental structure. *McGill Working Papers in Linguistics*, *9*, 180–210.

Cowan, N., & Morse, P. A.（1986）. The use of auditory and phonetic memory in vowel discrimination. *Journal of the Acoustical Society of America*, *79*（2）, 500–507.

Crowder, R. G.（1982）. Decay of auditory information in vowel discrimination. *Journal of Experimental Psychology: Human Learning and Memory*, *8*, 153–162.

Francis, A. L., & Ciocca, V.（2003）. Stimulus presentation order and the perception of lexical tones in Cantonese. *Journal of the Acoustical Society of America*, *114*（3）, 1611–1621.

Healy, A. F., & Repp, B. H.（1982）. Context independence and phonetic mediation in categorical perception. *Journal of Experimental Psychology: Human Perception & Performance*, *8*, 68–80.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B.（1992）. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*（5044）, 606–608.

Polka, L., & Bohn, O. S.（2003）. Asymmetries in vowel perception. *Speech Communication*, *41*, 221–231.

Repp, B. H.（1987）. On the possible role of auditory short-term adaptation in perception of the prevocalic［m］-［n］contrast. *Journal of the Acoustical Society of America*, *82*, 1525–1538.

Repp, B. H., & Crowder, R. G.（1990）. Stimulus order effects in vowel discrimination. *Journal of the*

Acoustical Society of America, *88* (5), 2080-2090.

Repp, B. H., Healy, A. F., & Crowder, R. G. (1979). Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance*, *5* (1), 129-145.

Repp, B. H., & Liberman, A. M. (1987). Phonetic boundaries are flexible. In S. R. Harnad (Ed.), *Categorical Perception: the Groundwork of Cognition* (pp. 88-112). New York, NY: Cambridge University Press.

Shigeno, S. (1992). Boin chikaku ni okeru bunmyaku kooka [The context effects in vowel perception]. *Onsei Gengo Igaku*, *33* (2), 204-208.

Tsushima, T. (1999). *Developmental changes in nonnative speech segment discrimination*. Unpublished Dissertation, Boston University, Boston, MA.

Tsushima, T. (2007a). Asymmetries in perception of an American English /r-l/ by adult Japanese learners of English. *Journal of the Japan Society for Speech Sciences*, *8*, 45-62.

Tsushima, T. (2007b). How did a stimulus order in AX discrimination training influence improvement in the ability to perceive an American English /l-r/ and /gl-gr/ among Japanese listeners? *Journal of the Acoustical Society of America*, *122* (5), 3017.

Tsushima, T. (2007c). Stimulus order effects in discrimination of the American English contrast /r-l/ by adult Japanese learners of English. *Journal of the University of Marketing and Distribution Sciences*, *19* (3), 15-27.

Tsushima, T., Sasaki, M., Shiraki, S., & Yoshida, K. (2001). Stimulus order effects in discrimination of non-native speech contrasts. *Journal of the University of Marketing and Distribution Sciences*, *13* (3), 17-35.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2003). On stimulus order effects in discrimination of nonnative consonant contrasts. *Acoustical Science and Technology*, *24* (6), 410-412.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2005). *Stimulus order effects in discrimination of a nonnative consonant contrast, English /b-v/, by Japanese listeners in the AX discrimination procedure*. Paper presented at the First Acoustical Society of America Workshop on L2 Speech Learning, Vancouver, Canada.

Tsushima, T., Yoshida, K., Shiraki, S., & Sasaki, M. (2003). How do stimulus variability and ISI influence stimulus order effects during discrimination training of nonnative consonant contrasts? *Journal of the Marketing and Distribution Sciences*, *15* (3), 73-89.

Tsushima, T., Yoshida, K., Shiraki, S., & Sasaki, M. (2005). Investigation into the mechanisms of the stimulus order effects in discrimination of nonnative speech contrasts. *Journal of the Marketing and Distribution Sciences*, *17* (3), 91-101.

Yamada, R., A. (1995). Age and acquisition of second language speech sounds: Perception of American English /r/ (inverted) and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 305-320). Timonium, MD: York Press.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: neural commitment to language-specific auditory patterns. *Neuroimage*, *26* (3), 703-720.