Ueda, R., & Hashimoto, K. (2019). Perceptual training in a classroom setting: Phonemic category formation by Japanese EFL learners. In J. Levis, C. Nagle, & E. Todey (Eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Ames, IA, September 2018 (pp. 237-249). Ames, IA: Iowa State University.

PRESENTATION/POSTER

PERCEPTUAL TRAINING IN A CLASSROOM SETTING: PHONEMIC CATEGORY FORMATION BY JAPANESE EFL LEARNERS

Ruri Ueda, Osaka Kyoiku University
Ken-ichi Hashimoto, Osaka Kyoiku University

Perceptual training targeting L2 phonemes has been reported as effective for both L2 learners' perception and production learning even without articulation practice. Considering the situation in EFL countries, especially Japan, where most English teaching and learning occur in classrooms with limited time, perceptual training can be an easy-to-conduct, effective method for L2 sound acquisition. Many of the studies reporting its positive effects, however, examined lab-based training, and only a few studies have tested the effects of perceptual training in a classroom setting. Therefore, to examine the applicability of perceptual training in the classroom, in the present study a ten-minute perceptual training targeting English /b/−/v/, /l/−/r/, and /s/−/θ/ was conducted in English courses at a university in Japan for six weeks. The results showed that students' scores on both the perception and production of /b/−/v/ and /s/−/θ/ significantly improved. However, the learning was not generalized to new word stimuli. For the /l/−/r/ contrast, neither their perception nor production performance changed after the training. Some possible reasons for smaller training effects than reported in many studies are discussed with reference to a lack of sufficient input and the way feedback was provided.

INTRODUCTION

For successful communication, it is necessary to perceive and produce the sounds of the language(s) used in communication. When it comes to L2 communication, it is widely agreed that learners' L1 has an influence on their L2 pronunciation performance (e.g., Tsukada, Birdsong, Bialystok, Mack, Sung, & Flege, 2005). Perception and production of L2 sounds absent from the learners' L1 sound system are said to be difficult because they do not have L2 phonetic representations or proper phonetic categories (e.g., Cutler, 2012; Flege, 1992). Several influential models have been proposed to explain the degree of difficulty in mastering each L2 phoneme according to the learner's L1, such as the Perceptual Assimilation Model (PAM; Best, 1995) and Flege's Speech Learning Model (SLM; 1995). According to PAM, the degree of difficulty perceiving each phoneme depends on how its phonemic contrasts are assimilated to learners' L1 phonemic categories. On the other hand, SLM explains that the greater the perceived dissimilarity of an L2 sound from the closest sound of the learner's L1, the more likely the acquisition of the L2 sound is. SLM also hypothesizes that perception and production share underlying representations, suggesting that improving perception skills by constructing phonetic representations can guide production learning as well.

Given the importance of skills to deal with L2 sounds in communication and the difficulty of mastering them, L2 educators have been concerned about effective ways of constructing L2 phonetic representations or proper L2 phonetic categories, building on the speech perception models described previously. One possible way is perceptual training, and many studies have reported its positive effects on learners' skills in perceiving and producing L2 sounds (e.g., Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005; Thomson, 2011). A typical format of perceptual training is an identification task with two-alternative forced choice (2AFC) format. In the task targeting the English /l/−/r/ contrast, for instance, participants first hear a stimulus like "lead." The minimal pair "lead" and "read" is then presented visually, and participants choose the word they think they heard. Immediate feedback is provided following the participant's choice. This type of perceptual training is considered effective for constructing L2 phonetic representations or modifying L2 phonetic categories because learners have exposure to extensive L2 sound input focusing on the target phonemes. More striking is that training could potentially lead to improvement of both L2 perception and production skills. For instance, Bradlow and her colleagues have extensively tested the influences of perceptual training on L2 learners' productions of the target sounds. In Bradlow et al. (1997), the participants were Japanese college students, and the target phonemic contrast was /l/ and /r/. The participants received perceptual training with the 2AFC format for 15 to 22.5 hours. The results showed that both their perception and production skills for the L2-English /l/−/r/ contrast improved. Also, the learning was generalized to non-familiar talkers and new words. A subsequent study reported that the learning effects persisted even three months after training (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999).

These findings suggest that perceptual training is effective, and that exposure to L2 sounds is important to improve L2 perception and production skills. However, L2 learners in EFL countries, such as Japan, have very limited opportunities to receive L2 sound input outside of the classroom. Given the importance of input in L2 acquisition (e.g., Thomson, 2011), it is ideal, and even essential, to provide sufficient sound input in class so that learners can acquire the skills to both perceive and produce sounds in the target language. We assume that perceptual training is an effective way to achieve this goal.

Most studies of the effects of perceptual training are lab-based, and only a few have reported on perceptual training in a classroom setting. In Hamada and Tsushima (2001), Japanese college students had three-week perceptual training on seven phonemic contrasts. The pre and posttest comparison found training effects on both perception and production skills. However, this study assigned out-of-class training sessions as homework as well as in-class training, which made it difficult to examine the effects of classroom-based training alone. Considering the possibility of applying perceptual training in the classroom, it is essential to examine whether the training effects observed in lab-based studies are also found in the classroom. Therefore, the present study examined the effects of classroom-based perceptual training for L1-Japanese learners on the pronunciation of L2-English consonants. The following research questions were addressed:

1. Does classroom-based perceptual training improve both the perception and production of L2-English phonemes?
2. Does classroom-based perceptual training generalize to accurate perception and production of new stimuli that did not appear during training?

METHODOLOGY

Participants

The study was conducted in two weekly English communication classes at a Japanese university for L1-Japanese college students. One class served as a treatment group (n = 24) and the other served as a control (n = 25). Only the data from the participants who attended every session during the study period were used in the analysis, leaving 13 students in each group. The overall English proficiency of the control group was slightly higher than that of the treatment group, which requires us to interpret results with caution.

Procedure

The study had a pre and posttest design. In week 1, both groups had a pretest session (see details below). From weeks 2 to 7, the treatment group had a weekly 10-minute perceptual training at the beginning of class, and the control group had regular English conversation practice instead. In week 8, both groups took a posttest session. In weeks 9 to 14, the content of the two groups was flipped in consideration of research ethics.

Training

The training took the format of a 2AFC task with the High Variability Training technique (Lively, Logan, & Pisoni, 1993), which consisted of stimuli spoken by four talkers (two female, two male). The target contrasts were /b/−/v/, /l/−/r/, and /s/−/θ/, which L1-Japanese speakers have difficulty in both perceiving and producing (Lambacher, 1999). Words used in the training were base−vase, berry−very, best−vest, bought−vote, lane−rain, late−rate, lead−read, lock−rock, seem−theme, sick−thick, sing−thing, and sum−thumb. The participants had 16 trials for each contrast each day of the training session, which means they had 48 trials in total per day.

The major difference of the present perceptual training from many reported studies was being conducted in class. The training video was projected on the screen at the front of the classroom so that all the students could do the task at the same time. The video contained a set of sound stimuli, a visual probe of the minimal pair, and immediate feedback for each training item, the last to facilitate the participants checking their answer immediately on their own. Each trial started with a beep, then the participants heard a word, after which they had to mark their answer on the answer sheet before a chime sounded. The correct answer was then zoomed and colored in red as feedback. The next trial then started with another beep.

Pre and posttests

The pre and posttests comprised two tasks, a 2AFC task to assess learners' perceptual skills and a word-list reading task for production skills. The former task was conducted as in the training session except that no feedback was provided in the tests. The pretest had 96 trials and the posttest had the same 96 trials and 24 additional trials to test learning generalization. For analysis, the posttest stimuli were divided into four conditions. Stimuli A were the same training stimuli spoken by two of the four speakers used in the training session; Stimuli B were the same training words spoken by non-familiar talkers; Stimuli C were new words spoken by two of the talkers used in the training session; and Stimuli D were new words spoken by non-familiar talkers. Conditions B, C, and D were established to examine the generalizability of the perceptual training.

In the word-list reading task, the participants were recorded reading aloud randomly-listed words from the posttest identification task by themselves. Later, the speech data were judged by three L1-English speakers with a 2AFC task format.

RESULTS

The perception skill results

Table 1 shows the perception scores for the training words spoken by familiar talkers in the pre and posttests.

Table 1

The identification task results of the training words spoken by familiar speakers

		Treatment group		Control group	
Contrast		Pre	Post	Pre	Post
/b/−/v/	*Mean*	10.85	13.38	12.46	12.85
	S.D.	2.51	1.61	3.13	1.99
/l/−/r/	*Mean*	8.62	8.85	9.31	9.69
	S.D.	2.75	2.48	2.25	2.78
/s/−/θ/	*Mean*	11.31	13.92	12.69	12.92
	S.D.	2.25	1.32	2.63	1.89
Totals	*Mean*	30.77	36.15	34.46	35.46
	S.D.	5.64	4.62	4.93	4.05

Overall, the treatment group significantly improved, while the control group did not. A series of two-way mixed ANOVAs with Time as a within-participant factor and Group as between-participant were conducted on the participants' identification scores of the /b/−/v/, /l/−/r/, and /s/−/θ/ contrasts. The alpha level was set at .017 to avoid Type I Error in multiple ANOVAs. For the /b/−/v/ contrast, the main effect of Time was observed [$F(1, 24) = 9.92$, $p < .004$, partial $\eta2 = .293$], while Group effect was not significant [$F(1, 24) = 0.44$, $p = .512$, n.s., partial $\eta2 = .018$]. The Time × Group interaction approached significance [$F(1, 24) = 5.39$, $p < .029$, partial $\eta2 = .183$]. Post hoc analyses using a Holm's Sequentially Rejective Bonferroni Procedure revealed a significant difference in the scores between the pre- and post-identification tasks in the treatment group [$F(1, 12) = 23.25$, $p < .001$, partial $\eta2 = .660$], but not in the control group [$F(1, 12) = 0.25$, $p < .624$, n.s., partial $\eta2 = .021$]. As for the /l/−/r/ contrast, none of the main effects nor their interaction was significant ($Fs < 1$). The results for the /s/−/θ/ contrast were almost identical to those for the /b/−/v/ contrast. The main effect for Time was significant [$F(1, 24) = 14.80$, $p < .001$, partial $\eta 2 = .381$], while the Group effect was not significant [$F(1, 24) = 0.07$, $p = .794$, n.s., partial $\eta 2 = .003$]. The Time × Group interaction was significant [$F(1, 24) = 10.39$, $p < .004$, partial $\eta 2 = .302$]. Post hoc analyses revealed a significant difference in only the treatment group [$F(1, 12) = 20.10$, $p < .001$, partial $\eta 2 = .626$], not in the control group [$F(1, 12) = 0.26$, $p < .621$, n.s., partial $\eta 2 = .021$].

Table 2 shows the posttest perception scores for the training words spoken by non-familiar talkers.

Table 2

The identification task results of the training words spoken by non-familiar speakers

		Treatment group	Control group
/b/−/v/	*Mean*	10.00	9.23
	S.D.	2.88	1.48
/l/−/r/	*Mean*	8.69	10.15
	S.D.	2.97	2.90
/s/−/θ/	*Mean*	14.38	12.69
	S.D.	1.08	1.81
Totals	*Mean*	33.08	32.08
	S.D.	5.43	4.97

Overall, the scores of the two groups were almost the same between the pre and posttests. An advantage for the treatment group was found in the /s/−/θ/ contrast, but not in the /b/−/v/ and /l/−/r/ contrasts.

Table 3 shows the posttest scores when the participants listened to new words.

Table 3

The identification task results of the new words spoken by familiar and non-familiar speakers

Contrast		Treatment group		Control group	
		Familiar	Non-familiar	Familiar	Non-familiar
/b/−/v/	*Mean*	2.62	2.46	2.92	2.92
	S.D.	0.74	0.63	0.73	0.73
/l/−/r/	*Mean*	1.92	2.00	2.38	3.00
	S.D.	1.14	1.18	1.08	1.04
/s/−/θ/	*Mean*	3.54	3.15	3.23	3.38
	S.D.	0.63	0.66	0.58	0.62
Totals	*Mean*	8.08	7.62	8.54	9.31
	S.D.	1.21	1.50	1.74	1.68

When the talkers were familiar, the control group's scores were better than the treatment group's overall. However, a series of two-way Group × Familiarity ANOVAs performed on each of the contrasts with an adjusted alpha level of .017 showed no significant main effects or interaction ($Fs < 4.38$). The trend was almost the same when the talkers were not familiar. This time, the control group was consistently better, but again, only numerically.

The production skill results

Table 4 shows the production results for the training words, which suggest that the treatment group showed a larger improvement than the control group.

Table 4

The production results for the training words

Contrast		Treatment group		Control group	
		Pre	Post	Pre	Post
/b/−/v/	*Mean*	12.92	16.15	16.08	17.23
	S.D.	2.10	4.95	3.64	4.64
/l/−/r/	*Mean*	13.85	14.69	13.31	14.31
	S.D.	3.72	5.42	4.33	4.64
/s/−/θ/	*Mean*	13.62	18.62	16.00	17.92
	S.D.	2.02	3.64	2.83	4.07
Totals	*Mean*	40.38	49.46	45.38	49.46
	S.D.	5.33	10.38	6.16	7.53

The scores for each contrast showed particularly large improvements for the /b/−/v/ and /s/−/θ/ contrasts in the treatment group, as suggested by two-way Time × Group mixed ANOVAs (alpha adjusted at .017 again). For the /b/−/v/ contrast, the main effect of Time was significant [$F(1, 24) = 7.54$, $p = .011$, partial $\eta 2 = .239$], while the Group effect was not [$F(1, 24) = 2.47$, $p = .129$, n.s., partial $\eta 2 = .093$]. The Time × Group interaction was not significant [$F(1, 24) = 1.69$, $p = .206$, n.s., partial $\eta 2 = .066$], though larger improvement was found in the treatment group. For the /l/−/r/ contrast, neither of the main effects nor their interaction was significant (Fs < 1.81). For the /s/−/θ/ contrast, the main effect of Time was significant [$F(1, 24) = 23.74$, $p < .001$, partial $\eta 2 = .497$]. Neither the Group effect [$F(1, 24) = 0.65$, $p = .429$, n.s., partial $\eta 2 = .026$] nor the Time × Group interaction [$F(1, 24) = 4.69$, $p = .041$, partial $\eta 2 = .164$] was significant. As the interaction effect approached significance, post hoc analyses were performed, revealing a significant difference in the average scores between the scores for the first and second recordings in both the treatment [$F(1, 12) = 30.95$, $p < .001$, partial $\eta 2 = .721$] and control groups [$F(1, 12) = 6.11$, $p = .021$, partial $\eta 2 = .203$].

Finally, Table 5 shows the production results for the new words.

Table 5

The production results for the new words

Contrast		Treatment group		Control group	
		Pre	Post	Pre	Post
/b/−/v/	*Mean*	4.00	4.23	4.77	4.46
	S.D.	1.41	1.42	1.36	1.61
/l/−/r/	*Mean*	3.62	3.00	3.00	3.31
	S.D.	1.04	1.96	1.63	1.63
/s/−/θ/	*Mean*	3.54	4.38	3.92	4.31
	S.D.	2.03	1.56	1.38	1.11
Totals	*Mean*	11.15	11.62	11.69	12.08
	S.D.	3.29	3.25	2.21	2.69

The same series of two-way mixed ANOVAs (Time × Group) with adjusted alpha of .017 were conducted. For the /b/−/v/ contrast, there were no significant main or interaction effects (Fs < 1.16), suggesting that neither group showed observable improvement. For the /l/−/r/ contrast, the main effects and their interaction were all non-significant (Fs < 3.18). Finally, for the /s/−/θ/ contrast, the treatment group showed improvement, while the control group did not, though the improvement did not reach significance (Fs < 2.60).

Correlation analyses

Another way to test the effectiveness of the perceptual training is to examine the correlations of the perception and production scores. Larger correlation coefficients should be observed for the scores for contrasts where the perceptual training had a positive effect. Table 6 shows the results of the correlation analysis of the perception and production scores for the pre and posttests. Curiously, we observed larger correlation coefficients on the posttest for all the contrasts in the treatment group, including the /l/−/r/ contrast that showed little training effect.

Table 6

The correlation coefficients of the perception and production scores

Contrast	Treatment group		Control group	
	pre	post	pre	post
/b/−/v/	-.19	.42	.47	.48
/l/−/r/	.12	.71	.39	.53
/s/−/θ/	.27	.60	.34	.09
Totals	.26	.69	.52	.57

To consider this further, we examined how the scores of individual participants changed. Figure 1 is a correlation plot of the overall scores for the treatment group, which mostly showed constant improvements in perception and production performance. This suggests that their phonetic representations of the target phonemes stabilized or their L2 phonetic categories were modified during training.

A similar trend was observed for the /s/−/θ/ contrast, as shown in Figure 2. Most of the participants improved in both perception and production, suggesting more stable representations or modified phonetic categories.

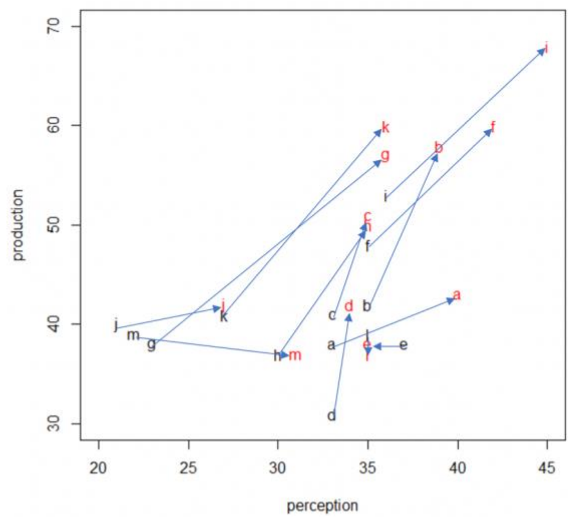

Figure 1. Transitions of overall perception and production scores by participants between the pre and posttests.

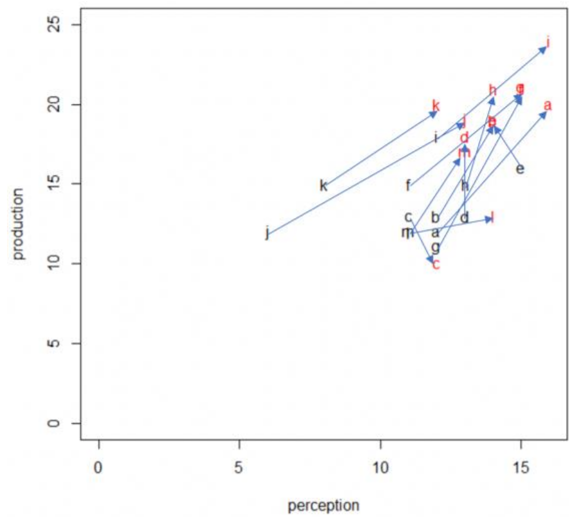

Figure 2. Transitions of /s/−/θ/ perception and production scores by participants between the pre and posttests.

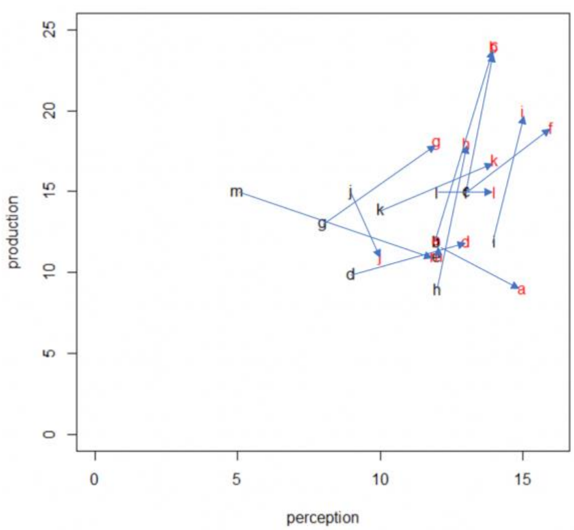

Figure 3. Transitions of /b/−/v/ perception and production scores by participants between the pre and posttests.

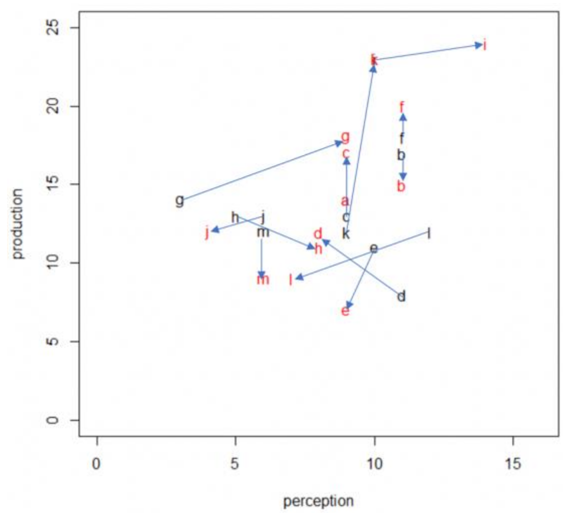

Figure 4. Transitions of /l/−/r/ perception and production scores by participants between the pre and posttest.

As for the /b/−/v/ contrast shown in Figure 3, again, we see constant improvement in perception, but the production scores changed more randomly. Some showed a dramatic improvement, while others were worse on the posttest. These somewhat mixed changes may have led to the weaker correlation on the posttest.

Finally, Figure 4 is a plot for the /l/−/r/ contrast, which showed no clear patterns. Although some participants improved in both perception and production, others showed a totally opposite trend, getting lower scores for both perception and production on the posttest. The posttest scores clustered around the regression line, which may have caused higher correlation coefficients on the posttest.

DISCUSSION

To sum up the results with reference to the research questions mentioned above, the answer to Research Question 1 is partially affirmative, with evidence that the present perceptual training improved the participants' perception and production of the familiar /b/−/v/ and /s/−/θ/, but not /l/−/r/ sounds. As for Research Question 2, however, the answer was negative, as the learning effect did not generalize to untrained stimuli except for the training words of the /s/−/θ/ contrast spoken by non-familiar talkers.

The training conducted in the present study was not effective in changing learners' /l/−/r/ perceptual and productive performances. One factor is the differences in how the three contrasts are assimilated to learners' L1 categories based on the PAM (Best, 1995). The English /l/−/r/ contrast is categorized as "single-category assimilation" by Japanese L1 learners, which means that both sounds are assimilated to the same L1 category to the same extent. On the other hand, the English /b/−/v/ and /s/−/θ/ contrasts are categorized as showing "category-goodness difference," that is, "Both sounds are assimilated to the same L1 category, but one is a far better match to it than the other" (Cutler, 2012, p. 306). According to the PAM, "single-category assimilation" is more difficult to learn than "category-goodness difference," making /l/−/r/ more difficult to acquire than /v/ and /θ/ for the present participants.

In addition, differences in the manner of articulation might also have affected the learning difficulties for each contrast. According to a meta-analysis by Sakai and Moorman (2018), perceptual training was more effective on obstruents than sonorants. They suggested that because obstruent sounds are articulated more saliently, learners can perceive the differences in sound more easily, which facilitates their learning new phonemes. In the present study, the English /b/−/v/ and /s/−/θ/ contrasts involve obstruents, while the /l/−/r/ contrast sonorants. In particular, discriminating English /l/−/r/ requires detecting formant differences. However, L1-Japanese speakers have difficulty in utilizing formant information to discriminate English /l/−/r/. For example, although F3 frequency plays an important role in discriminating English /l/−/r/ sounds, L1-Japanese speakers tend to rely on F2 frequency, which is insufficiently reliable in /l/−/r/ discrimination (Iverson, Hazan, & Bannister, 2005). Therefore, catching the differences underlying the /l/−/r/ contrast should have been more difficult for the present participants.

Regarding the learning generalization, the training effects were not generalized to new stimuli in either perception or production. In studies reporting learning generalization, the participants had 15–22.5 hours of training for each pair of target phonemes, while a recent study of Qian, Chukharev-Hudilainen, and Levis (2018) reported that learning was not generalized to new words when the participants had only 10- to 100-minute training for 12 contrasts. Because the training conducted in the present study lasted only an hour for six phonemes, the lack of sufficient input might be the primary reason why the training effect failed to generalize.

Another possible reason for the lack of learning generalization is the way feedback was provided. Most of the studies that reported positive effects of perceptual training were lab-based. However, the training conducted in the present study was classroom-based. A major difference between the two is how learners receive feedback. In lab-based training, learners receive individualized feedback, while in classroom-based training, answers are presented on the screen item by item, which the learners need to check by themselves. In such a situation, inattentive learners can easily miss the correct answer and do not notice whether or not they made a mistake. Previous studies showed that corrective feedback facilitates improvement of L2 speech perception (Lee & Lyster, 2015) and production (Lyster, Saito, & Sato, 2013), since feedback gives L2 learners opportunities to modify their knowledge. Failure to utilize feedback information effectively might be another source of the decreased learning effects in the present study.

CONCLUSION AND FUTURE STUDIES

The present study examined the effects of perceptual training of L2-English phonetic contrasts in a classroom setting for L1-Japanese college students in Japan. The training led to improvements in the /b/−/v/ and /s/−/θ/ contrasts but not the /l/−/r/ contrast in both perception and production. Moreover, while the training effects were generalized to the training words spoken by non-familiar talkers for the /s/−/θ/ contrast, none of the contrasts were generalized to new words.

There are two issues that need to be addressed in the future studies. First, as mentioned in the discussion, the amount of input is critical. Therefore, we would like to conduct perceptual training with more input to see if learners' /l/−/r/ performances improve. The other critical issue concerns feedback. In the present study, learners were simply shown the correct answer for each item on the screen. In a future study, learners will be asked to mark their answers by themselves to elicit greater attention to the correctness of their answers to see whether this leads to more effective learning. Eventually, in future studies we would like to examine the proposed models of L2 phonetic acquisition and the relations between L2 sound perception and production.

ABOUT THE AUTHORS

Ruri Ueda, M.Ed. (ruri.u10@gmail.com), is a research student in the Department of English Education at Osaka Kyoiku University. She has researched the effects of classroom-based perceptual training and flipped classrooms. Also, she presented a study of differences in L2 learners' pronunciation performances by degree of task spontaneity at the 9th PSLLT Conference. She is currently interested in the relationship between L2 perception and production at the segmental level.

Ken-ichi Hashimoto, Ph.D. (kenhashi@cc.osaka-kyoiku.ac.jp), is an associate professor in the Department of English Education at Osaka Kyoiku University. His research interests lie in the psycholinguistics of second language acquisition, and he obtained his PhD degree from the University of Queensland, Australia, with a dissertation on the effects of L1 syntactic properties on L2 relative clause processing. Since his PhD study, he has been interested in various factors affecting L2 sentence comprehension processes, including L1 influences and interactions between lower-level (e.g., phonetic and lexical) and higher-level (e.g., syntactic) processes.

Contact information for both authors:
Osaka Kyoiku University
4-698-1, Asahigaoka, Kashiwara,
Osaka, 582-8582, Japan
+81-72-978-3525

REFERENCES

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167–200). Timonium, MD: York Press.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics, 61*, 977–985. doi:10.3758/BF03206911

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*(4), 2299–2310. doi:10.1121/1.418276

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge: MIT Press.

Flege, J. E. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In R. D. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement, and management* (pp. 157–232). Amsterdam: John Benjamins. doi:10.1075/sspcl.1.06fle

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Hamada, M., & Tsushima, T. (2001). Relation between improvements of perception and production ability during a university English speech training course. In J. White (Ed.), *Proceedings of the Fourth Conference on Foreign Language Education and Technology* (pp. 579–582). Kobe: Japan Association for Language Education and Technology.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/−/l/ to Japanese adults. *The Journal of the Acoustical Society of America, 118*(5), 3267–3278. doi:10.1121/1.2062307

Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning,12*(2), 137–156. doi:10.1076/call.12.2.137.5722

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*(2), 227–247. doi:10.1017/S0142716405050150

Lee, A. H., & Lyster, R. (2016). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition, 38*(1), 35–64. doi:10.1017/S0272263115000194

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242–1255. doi:10.1121/1.408177

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching, 46*(1), 1–40. doi:10.1017/S0261444812000365

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning & Technology, 22*(1), 69–96. doi:10125/44582

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics, 39*(1), 187–224. doi:10.1017/S0142716417000418

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowels: Perception improves pronunciation. *CALICO Journal, 28*(3), 744–765. doi:10.11139/cj.28.3.744-765

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics, 33*(3), 263–290. doi:10.1016/j.wocn.2004.10.002