Effects of Corrective Feedback on Early and Late L2 Learners' Perception of /θ/-/s/

Ying Li[1,*]

[1]School of Foreign Languages, Southwest University of Political Science and Law, China

*Correspondence: School of Foreign Languages, Southwest University of Political Science and Law, China. E-mail: liying_22@163.com

Received: March 7, 2016 Accepted: March 15, 2016 Online Published: March 20, 2016

doi:10.5430/wjel.v6n1p14 URL: http://dx.doi.org/10.5430/wjel.v6n1p14

Abstract

This study explores whether both early and late second language (L2) learners are able to benefit from instruction that includes corrective feedback (CF) on the perception of L2 speech sounds. It is a classroom-based study. Four groups of participants with less than 60% accuracy in the perception of /θ/-/s/ were recruited. They were all Mandarin speakers of English. Group 1 (n=30, mean age: 21.23) and group 2 (n=16; mean age=20.31) began L2 English study at or before 7 years old, whilst group 3 (n=29; mean age=19.48) and group 4 (n=17; mean age=20.56) started L2 English learning at or after the age of 12. Group 1 and group 3 went through a four-week programme of form-focused lessons that drew their attention to the non-native contrast /θ/-/s/; they were given feedback during the awareness task. Group 2 and group 4 went through another four weeks of form-focused lessons with the same instruction and awareness tasks, except that no feedback was given during the awareness tasks. All the participants' accuracy regarding the perception of /θ/-/s/ was tested before the lessons (pre-test), immediately after the lessons (post-test), and one month after the lessons (delayed test). The results show that (1) compared with the pre-test, the four groups all received some degree of improvement in the post-test and delayed test, yet group 1 and group 3 displayed significantly higher accuracy than group 2 and group 4; (2) unexpectedly, the difference between group 1 and group 3 as well as the difference between group 2 and group 4 was statistically insignificant. The findings suggested that instruction with corrective feedback could be effective both in facilitating early and late learners' perception of non-native speech sounds. The effect of onset age on L2 learning, however, may not be overstated in L2 speech learning.

Keywords: *corrective feedback; speech perception; L2; onset age*

1. Instruction

Research on second language (L2) acquisition has shown us that there are various factors affecting language learners' acquisition of non-native speech sounds, such as the influence of L1 knowledge, motivational factors, etc. (Oyama, 1976; Best, 1995a, b). One of the long-standing topics of debates is the age factor. No consensus has been achieved regarding whether late language learners are able to achieve native-like proficiency in an L2 (Bohn & Flege, 1990; Escudero & Boersma, 2004; Flege, Bohn, & Jang, 1997). Another factor that has attracted our attention is the role of corrective feedback in L2 acquisition, which is gaining increasing prominence as scholars have looked specifically into its role in different fields of L2 learning (Doughty & Varela, 1998; Havranek, 1999; Lyster & Ranta, 1997; Ohta, 2000; Oliver, 2000). In the field of L2 speech learning, unfortunately, the majority of the existent studies were carried out in laboratories instead of real classrooms, which simply gave the participants feedback in terms of whether their answers were true or false without any specific explanation or further instruction (i.e., Li, 2015; Hardison, 2003; Logan, Lively, & Pisoni, 1991). This study therefore examines the role of corrective feedback in early and late language learners' acquisition of an L2 contrast. Specifically, it focuses on the participants' perception of the target contrast. What makes the present study different from previous ones is that it was carried out in real classrooms.

1.1 Age Factor

The age factor has long been considered as a significant variable affecting the degree of achievement in L2 learning.

Traditional views such as the Critical Period Hypothesis (CPH) claim that L2 learners are unable to achieve native-like proficiency if they commence L2 study after the end of a "critical period" (Lenneberg, 1967), or "sensitive period" (Oyama, 1976), which is usually defined as the period of puberty; its assumption rests on the fact that as a language learner's brain matures, its plasticity in terms of language acquisition is gradually lost. There are some findings in support of CPH. For example, in Mayo, Florentine, & Buus (1997), the perception of English words in noise for native English speakers, early bilinguals and late bilinguals was tested and compared. The early bilinguals were native Mexican-Spanish-speaking listeners who learned English before the age of six. The late bilinguals were Mexican-Spanish-speaking listeners who learned English after the age of 14. The results show that the native English speakers and early bilinguals displayed significantly higher intelligibility and benefited significantly more from the context than the late learners. Another supporting example comes from Shi (2010); in this study, native bilingual, and early, late, and very late non-native bilingual learners' perception performance was tested. The participants were asked to identify target words in 400 Speech-Perception-in-Noise sentences presented in different combinations of noise and context. The findings indicate that the age of L2 acquisition had a significant effect on their perception performance. Specifically, the very-later learners' performance was found to be significantly compromised in all test conditions.

However, CPH also suffered criticism from the perspectives of theoretical models (i.e. Flege's Speech Learning Model) and experimental findings (Fullana and Mora, 2008; Flege, Munro, & MacKay, 1995; Yamada, 1995). For instance, Flege's Speech Learning Model (SLM) (Flege, 1981, 1987, 1988, 1991, 1992a, b, 1995a, b, 2003) examines the constraints of L1 experience on the acquisition of L2 speech sounds for L2 learners, proposing an extensive set of hypotheses on this issue. One of the most significant assumptions is that language learners' capacity in speech learning remains intact throughout their life, which directly opposes CPH. SLM attaches great importance to the role of input in L2 acquisition. According to SLM, with sufficient input, L2 learners are eventually able to acquire a non-native speech sound by creating a new category of the target sound, despite their OA differences. This view has been illustrated by numerous investigations of speech perception such as Hazan, Sennema, & Faulkner (2005) and Iverson & Evans (2009). In these studies, adult language learners were successfully trained to perceive and/or produce the target non-native speech sounds.

Another model which provides counterevidence to CPH is Best & Tyler's (2007) Perception Assimilation Model-L2 (PAM-L2). According to PAM-L2, a learner's L1 has a significant effect on their perception of L2 speech sounds. In fact, PAM-L2 argues that child learners have an advantage over adult learners in the perception of L2 sounds. The reason, however, is because child learners have comparatively less L1 experience than adults, which interferes with their perception of L2 sounds. Nonetheless, PAM-L2 suggests that adult learners are neither uniformly poor at the perception of all L2 sounds, nor incorrigible in identifying the L2 sounds that they initially have difficulty with (Best, 1994). L2 listeners' level of success in perceiving an L2 speech sound is largely dependent on the assimilation of the L2 sound with its counterpart in the listeners' L1. Specifically, PAM-L2 predicts four possible types of assimilation that may occur. In the first situation, listeners tend to equate an L1 sound to a correlated L2 sound on a phonological level, despite their phonetic difference(s). For example, native English listeners are likely to equate the French /r/ with the English /ɹ/. In the second situation two L2 sounds may be perceived as the same phonetic category, yet one is perceived as a better exemplar than the other. Third, two L2 speech sounds are assimilated into a single L1 phonetic category, which is the most difficult situation for listeners. The last situation occurs when an L2 sound can be perceived without being assimilated to the listener's L1 category, meaning no assimilation occurs. Regardless of which type of assimilation takes place, PAM-L2 predicts that even adult L2 learners are able to learn to perceive L2 sounds with sufficient input (Best, 1994, 1995a, b).

In the aspect of experimental studies, we can also find counterevidence to CPH. For instance, in Li's (2015) experiment, 29 adult L1-Chinese of L2-English speakers were exposed to nine sessions of audio-visual perception training with identification tasks being carried out. Their perception and production of the target contrasts /θ/-/s/ and /ð/-/z/ was tested before, during, and after the training sessions. According to the results, compared with the pre-test the 29 participants' accuracy in the perception and production of the target contrasts improved significantly in the mid-test and post-test. More importantly, none of the participants had started L2-English learning before 12 years old. Similar findings are also available from Fullana & Mora (2008) and Yamada (1995).

Due to the inconsistent findings from previous studies, we still cannot be sure whether language learners of younger OAs are better able to perceive non-native speech sounds than those with older OAs.

1.2 Corrective Feedback

In the field of L2 acquisition, corrective feedback is frequently defined as "any feedback provided to a learners, from

any source, that contains evidence of learner error of language form", which can either be "oral or written, implicit or explicit" (Russell & Spada, 2006). Some laboratory speech perception training programmes carried out by former scholars suggest that phonetic training with corrective feedback could facilitate acquisition of the target speech sounds, for example Logan, Lively, & Pisoni (1991), Hardison (2003), and Li (2015). In these studies, the significance of immediate feedback was emphasized, because it enables learners to consistently focus on a key feature of a target sound (Lee & Lyster, 2015) and helps the participants to maintain and increase their attention during the tasks (McGuire, 2010). Given that L2 learners are more likely to be exposed to real classroom settings than in laboratories (Lee & Lyster, 2015), it would be necessary to investigate the effects of an instructional input with corrective feedback on the acquisition of L2 sounds.

Unfortunately, there is a paucity of evidence regarding the effects of corrective feedback on L2 speech perception in classroom settings. Recently, Lee & Lyster (2015) conducted a classroom-based study which examined its effects on learners' perception of L2 speech sounds; 32 adult L1-Korean of L2-English learners were exposed to five one-hour form-focused lessons that drew their attention to English /i/-/ɪ/. Some of the participants were given oral corrective feedback in identification tasks carried out in the lessons (instruction+CP group), whereas others were not (instruction-only group). The instruction+CP group outperformed the instruction-only group in the post-test and the delayed post-test. Due to lack of further evidence on this issue, it is still unclear whether instruction with corrective feedback is better able to facilitate L2 perception compared to instruction without corrective feedback. The present study therefore aims to further shed some light on this issue.

1.3 Linguistic Targets: /θ/-/s/

The target contrast of the present study was /θ/-/s/. The participants were L1-Mandarin of L2-English speakers. As discussed by Li (2015), English /θ/ and /s/ are different in terms of articulatory characters and acoustic features. Specifically, English /θ/ is inter-dental fricative, whereas English /s/ is alveolar fricative. As for acoustic features, English /θ/ shows a lower frequency range, amplitude range, inherent duration, and relative intensity compared to /s/; /θ/ also displays a vowel transition in terms of downward F2, while /s/ does not show the vowel transitional feature. The comparison of Mandarin and English phonetic inventory shows us that Mandarin /s/ is the same as, or very similar to English /s/ in terms of articulatory and acoustic features, depending on speaker differences[1]. English /θ/, however, does not exist in Mandarin. Considering the influence of L1 on the acquisition of L2 (i. e., Flege's SLM; Best & Tyler's PAM), it is thus speculated that Mandarin learners of English may have difficulty in the acquisition of English /θ/. Some previous studies actually have illustrated this speculation. For example, in Li's (2015) study, the university-level participants with intermediate English proficiency were found to have serious difficulties in distinguishing /θ/ from /s/.

2. Methodology

2.1 Participants

Questionnaires (see Appendix I) were handed out at a University in Chongqing China to select potential participants with different OAs of L2-English learning. As a result, 102 students who started L2-English learning before the age of seven and another 143 students whose OA of L2-English learning was over 12 were selected to take part in a pre-test. In the pre-test, the 245 participants' accuracy in the perception of /θ/-/s/ was tested. In accordance with to the results, 92 participants with perception accuracy below 60% were recruited to join the following sessions of the study. As shown in Table 1, they were divided into four groups: participants in group 1 (n=30 (16 female, 14 male); mean age: 21.23) and group 2 (n=16 (8 female, 8 male); mean age=20.31) began L2 English study before the age of seven; group 3 (n=29 (15 female, 14 male); mean age=19.48) and group 4 (n=17 (9 female, 7 male); mean age=20.56) started L2 English learning after the age of 12. All the participants were L1-Mandarin of L2-English learners and were non-English major undergraduates. They were reported to have normal hearing capacity.

Table 1. Age and OA Information for the Four Groups

	Group 1		Group 2		Group 3		Group 4	
N	30		16		29		17	
	OA	Age	OA	Age	OA	Age	OA	Age
Mean	6.10	18.77	12.56	18.50	6.28	18.34	12.47	18.47
Range	3	3	4	2	3	3	2	3
Minimum	4	17	10	18	4	17	12	17
Maximum	7	20	14	20	7	20	14	20

2.2 Stimuli

The stimuli used in the pre-test and post-test were the same as those employed in Li (2015) (see Appendix II). Specifically, /θ/-/s/ were embedded in the initial, medial, and final positions of nonsense words with /i, a, u/ as the vowel context. The syllable structures were VC, VCV and CV^2. Nonsense words were used to avoid the influence of lexical knowledge on the participants' perception performance (Hazan et al., 2005). Each stimulus was produced twice by a female (27 years old) and a male (45 years old) received pronunciation speaker. They were both native English speakers and were phonetically trained. Their production was auditorily recorded with a high quality recorder (Roland-05) in a quiet room. The settings of the recorder were a 16-bit mono channel and 44.1 KHz for sampling frequency. The recordings were transferred to a laptop, and were then cut, repeated three times, and randomized. A total of 108 stimuli was yielded (*interstimulus interval*=1,000 ms). The stimuli were then coded into an AXB task with a script, which was carried out using Praat (Boersma & Weenink, 2015).

The stimulus words employed in awareness tasks included minimal pairs of both real and nonsense words. It was because there are limited number of words that contain /θ/ and /s/ as minimal pairs. The stimuli were employed from that used in Li (2015) for phonetic training.

2.3 Procedure

A university English teacher, who was a native Mandarin speaker but had been studying and teaching in the UK for 21 years since she was 13 years old, was asked to teach /θ/-/s/ to the participants. She was told the purpose of the study and was asked to clarify the two sounds as much as possible. As shown in Table 2 below, the instruction lessons took place four times a week for four weeks. Each time, the lesson lasted for about one hour. The four groups of participants were instructed separately by the same instructor at different times. As shown in Table 1, in each lesson the instructor began with the instruction of /θ/-/s/ pronunciation, particularly /θ/, in terms of articulatory gestures, the way air flows, etc. The instruction was carried out in Mandarin to ensure clear understanding. This was then followed by the instruction of /θ/-/s/ in different vowel/consonant contexts, and an awareness task—a pick-up-a-card game—was carried out for further practice. During the awareness task, groups 1 and 3 were given corrective feedback, whereas no feedback was given to groups 2 and 4.

Table 2. Schedule of the Instruction Lessons

week 1	week 2	week 3	week 4
lesson 1	lesson 1	lesson 1	lesson 1
1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /i/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /a/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /u/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in random consonant context 3. Awareness task
lesson 2	lesson 2	lesson 2	lesson 2
1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /æ/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /ʌ/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /au/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in random consonant context 3. Awareness task
lesson 3	lesson 3	lesson 3	lesson 3
1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /ɔ/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /e/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in /ai/ context 3. Awareness task	1. Pronunciation-focused instruction of /θ/-/s/ 2. Pronunciation-focused instruction of /θ/-/s/ in random consonant context 3. Awareness task

The post-test was conducted on the second day after the instruction lessons were completed. Each participant was asked to sit in front of a desktop that was equipped with a high-quality headphone. They were instructed to adjust the volume if they needed to, and they were asked to do an AXB task; among the three "words" they heard in each trial, they needed to choose whether X (the second one) was the same as A (the first one) or B (the second one) by mouse clicking on the corresponding button on the screen. A subsequent trial was triggered by clicking on an answer, and they were allowed to hear a trial as many times as they needed by clicking on the button "listen again". The participants' responding time and answers were automatically recorded with Praat (Boersma & Weenink, 2015). The delayed test was the same as the post-test, yet the stimuli were randomized and presented with different orders. This was carried out one month after the post-test.

2.4 Feedback

Group 1 and group 3 were given feedback in the awareness tasks. As there was only one instructor and there were about 30 participants in each of the groups, it was not easy to give oral feedback to individual participants during the class. The instructor provided them with immediate oral feedback of the right answer for each task after they responded. In the meantime, she encouraged the participants with words/sentences like "good," "perfect", "right", and "well done". Sometimes the instructor used a rising tone if any participants gave the wrong response in order to encourage self-repair. In each trial, after the right answer was given, the instructor asked the students to read the target word twice after her. This was done because speech perception and production are found to be closely related (Williams & McReynolds, 1975; Jamieson & Rvachew, 1992; Watkins, Strafella, & Paus, 2003), though there is no consensus on whether speech proceeds production or vice versa (Best & Tyler, 2007; Flege, 1991). It was hoped that the demonstration of the production of the words could further facilitate correct perception of the target sounds. As this study was limited in the aspect of speech perception, the participants' production errors, if any, were not corrected. Here is an excerpt from the pick-up-a card game.

(1) Instructor: Could you show me *sick*?

Students: [Show the instructor the card *sick* or *thick*]

Instructor: *thick*? (show the students the card of *thick*)

Students: Yes/No (depending on their own answers)

Instructor: Good! It is *sick*! (Show the students the card of sick)—Corrective feedback.

Instructor: Now read after me, *sick*, *sick*.

Students: *Sick*, *sick*.

3. Results

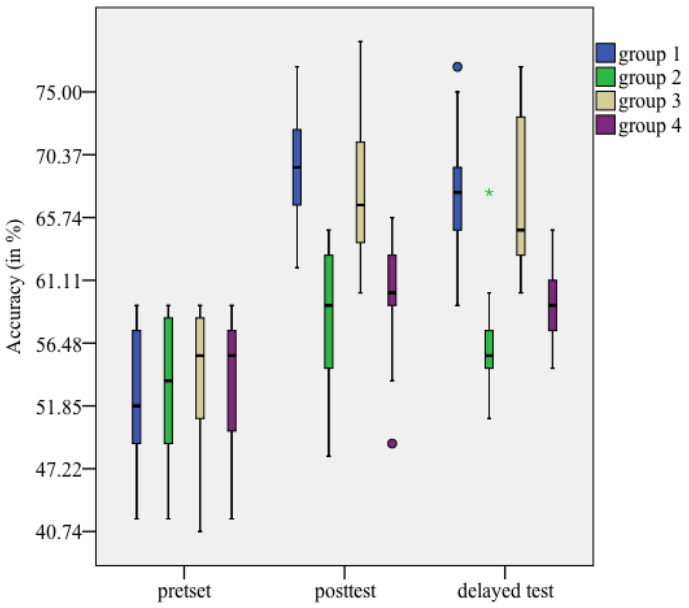

Figure 1. Boxplots of the Perception Accuracy in the Pre-Test, Post-Test, and Delayed Test for the Four Groups

Table 3. Descriptive Data of the Participants' Accuracy in the Pre-Test, Post-Test and Delayed Test (in %)

Tests	Group	N	Minimum	Maximum	Mean	Std. Error
Pre-test	group 1	30	43.52	59.26	52.72	0.95
	group 2	16	43.52	59.26	53.07	1.39
	group 3	29	40.74	59.26	53.55	1.09
	group 4	17	43.52	59.26	53.32	1.29
Post-test	group 1	30	62.04	76.85	69.35	0.72
	group 2	16	48.15	64.81	58.28	1.23
	group 3	28	60.19	78.70	68.22	1.06
	group 4	17	49.07	65.74	59.86	0.99
Delayed test	group 1	30	59.26	76.85	67.44	0.80
	group 2	16	50.93	67.59	56.31	0.96
	group 3	29	60.19	76.85	67.18	0.98
	group 4	17	50.93	64.81	58.12	0.74

Figure 1 and Table 4 show the four groups of participants' perception accuracy in the pre-test, post-test and delayed test. The figures are based on the percentages of correctness in the AXB tests. The participants' perception accuracy was the baseline of the study. As mentioned in section 2.1, only those participants whose accuracy was lower than 60% were recruited; thus, the maximum accuracy of the selected participants was 59.26%. The majority of their accuracy was between 40% and 59%. After the instruction lessons, the figure changed dramatically. The mean accuracy for groups 1 and 3 was about 10% higher than that of group 2 and group 4. Moreover, group 1 and group 3 retained a similar accuracy in the delayed test that was carried out one month after the instruction lessons.

Given the fact that the perception test was carried out with AXB tasks in which there were two possible responses, it was important to correct any potential bias in the participants' responses (Hazan et al., 2005). Therefore, the participants' accuracy was converted to *d-prime* scores[3] (*d'* scores thereafter) for further statistic analysis. Specifically, individual participant responses in each trial were used as inputs for SPSS. The hit-rate and false-alarm-rate were computed with Crosstabs first, and these were then converted into *d'* scores using Excel. The calculated *d'* scores were then input into SPSS for further statistic analysis.

Table 4. Descriptive Data of the Participants' *d'* Scores in the Pre-Test, Post-Test and Delayed Test

Test	Group	N	Minimum	Maximum	Mean	Std. Error
Pre-test	group 1	30	0.10	0.86	0.44	0.89
	group 2	16	0.10	0.84	0.46	1.06
	group 3	29	0.10	0.84	0.51	0.85
	group 4	17	0.10	0.84	0.47	1.16
Post-test	group 1	30	1.04	2.26	1.41	0.64
	group 2	16	0.17	1.17	0.82	1.02
	group 3	29	0.51	2.87	1.43	0.96
	group 4	17	0.33	1.27	0.86	0.75
Delayed test	group 1	30	0.71	2.26	1.29	0.63
	group 2	16	0.33	1.37	0.66	0.82
	group 3	29	0.93	2.26	1.33	0.77
	group 4	17	0.49	1.15	0.80	0.52

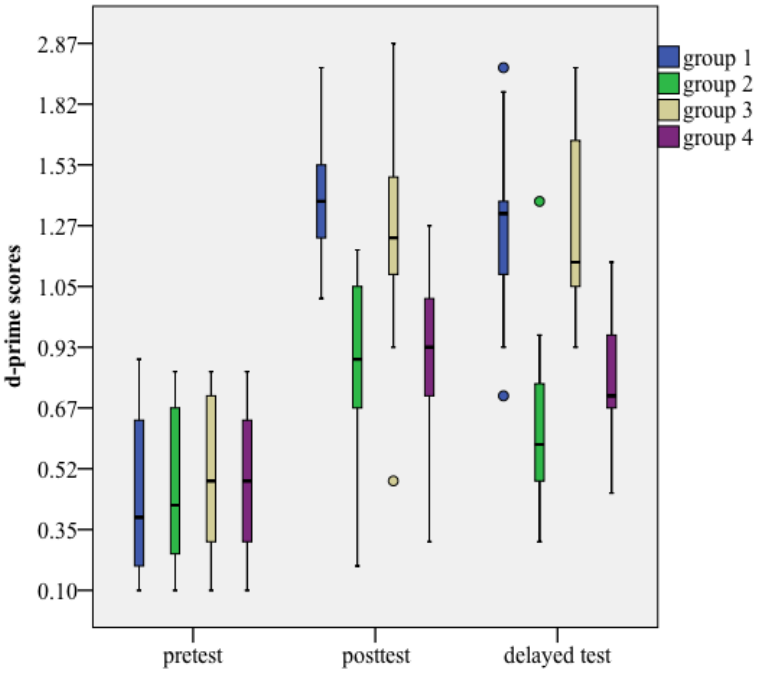

Figure 2. Boxplots of the Four Groups of Participants' D' Scores in the Pre-Test, Post-Test, and Delayed Test

Table 4 and Figure 2 display the participants' *d'* pre-test, post-test and delayed test scores. Similar to in the results shown in Table 3 and Figure 1 above, the four groups of participants displayed similar accuracy in the pre-test. *ANOVA* indicated that compared with the pre-test, the four groups' *d'* scores were significantly improved in the post-test and delayed test (group 1: $F(2, 89)=112.22$, $p<0.001$; group 2: $F(2, 47)=7.63$, $p=0.001$; group 3: $F(2, 86)=41.45$, $p<0.001$; group 4: $F(2, 50)=13.89$, $p<0.001$). Moreover, there were no significant differences among the *d'* scores in the pre-test ($F(3, 91)=0.47$, $p=0.71$). However, groups 1 and 3 displayed significantly higher accuracy than groups 2 and 4 in the post-test ($F(3, 91)=14.95$, $p<0.001$) and delayed test ($F(3, 91)=25.07$, $p<0.001$). The statistic difference between group 1 and group 3 was found to be insignificant in the post-test ($F(1, 58)=0.02$, $p=0.88$) and delayed test ($F(1, 58)=0.19$, $p=0.67$). Similarly, the difference between the d' scores of groups 2 and 4 were revealed to be statistically insignificant both in the post-test ($F(1, 32)=3.12$, $p=0.09$) and delayed test ($F(1, 32)=0.13$, $p=0.72$). Furthermore, neither age nor gender difference were significant factors in the participants' post-test or delayed test perception performance ($p>0.05$). The rest of the variables that were investigated with questionnaires, such as the length of English study, primary motivation in English study, whether traveled to or lived in English speaking countries, daily amount of English use, whether study English at spare time, were all found to have no significant effect on the participants' perception accuracy ($p>0.05$).

Considering that one aspect of the present study is to investigate whether learners' OA of L2 learning affects their perception learning outcomes, it is necessary to examine whether OA has played a significant effect on the participants' performance in the post-test and delayed test. Therefore, the difference in the participants' L2 learning OA was coded as the independent variable, while their *d'* scores in the post-test and delayed test were coded as dependent variables. According to the results, OA was revealed to be an insignificant variable in terms of the participants' perception performance in the post-test ($F(1, 91)=0.01$, $p=0.91$) and delayed test ($F(1, 91)=0.81$, $p=0.37$).

Moreover, in order to detect whether instruction lessons with/without corrective feedback played a significant role in the participants' perception accuracy, *one-way ANOVA* was conducted. Specifically, the participants' *d'* scores in the post-test and delayed test were coded as the dependent variable, while whether they received corrective feedback during the awareness tasks was coded as the independent variable. The results show that corrective feedback had a significant effect on their perception accuracy in the post-test ($F(1, 91)=45.72$, $p<0.001$) and delayed test ($F(1, 91)=71.19$, $p<0.001$).

4. Discussion

One of the major findings of the present study was that, compared with the pre-test, both early (group 1 and group 2) and late (group 3 and group 4) L2 English learners' accuracy was significantly improved after they went through the instruction lessons. Moreover, their improved perception ability remained one month after the post-test. This finding is at odds with the traditional view of CPH regarding the role of OA in L2 acquisition, which claims that L2 learners are unable to achieve native-like proficiency if they commence L2 study after the end of the "critical period" (Lenneberg, 1967). While it was true that none of the participants achieved 100% accuracy in the perception task, which might be achieved by native English speakers, the instruction lessons were only carried out for four weeks. With the significant improvement observed, we might be able to predict that the participants could receive higher perception accuracy with further instruction lessons. Moreover, there was no significant difference was revealed between the accuracy of group 1 and group 3, as well as group 2 and group 4. In other words, OA did not play a significant role in their perception performance. This finding is consistent with some previous studies, such as Hazan, Sennema, & Faulkner (2005) and Iverson & Evans (2009). Nevertheless, it seemed to be opposite to the findings of Mayo et al. (1997) and Shi (2010), in which the early bilingual beginners had better perception performance than the late bilingual beginners. The discrepancy might be caused by the fact that Mayo et al. (1997) and Shi (2010) only tested the early and late bilinguals' perception performance, whereas in the present study the participants went through instruction lessons that facilitated their perception accuracy.

Compared with some laboratory-based studies carried out by previous scholars (i e., Hazan et al., 2005; Iverson & Evans, 2009), the present study received similar success. This finding supports Flege's SLM and Best & Tayler's PAM-L2, both of which predict that even adult L2 learners are able to acquire non-native speech sounds which they initially have difficulty with. Instead of emphasizing the role of OA in L2 study, SLM and PAM-L2 both attached great significance to input in L2 acquisition. Given that the instruction lessons were revealed to have had a significant effect on perception improvement, this finding may confirm the critical role of input in L2 speech perception.

Moreover, PAM-L2 predicted four possible types of assimilation that may occur in L2 speech learning. As discussed in section 1.3, English /θ/ does not exist in the Mandarin phonetic system, whereas English /s/ is very similar to Mandarin /s/ in terms of articulatory gestures and acoustic features. Due to the lack of further acoustic tests and analysis, it was unclear which type of assimilation occurred among the participants. Nonetheless, in Li's (2016) study, the L1-Mandarin of L2-English speakers were found to have replaced /θ/ with /s/ in the production test. Therefore, it is possible that the selected participants, who showed less than 60% accuracy in the discrimination of /θ/-/s/, may have equated /θ/ with /s/. This type of assimilation was predicted to be difficult for the language learners in terms of learning the non-native sound (Best & Tyler, 2007), however the improved perception of accuracy suggests that even in this situation L2 learners are able to learn to correctly perceive the target non-native sound with instructional input.

The second significant finding of the present study is that the participants who were given feedback in awareness tasks displayed significantly higher improvement than those who were not given feedback. Some previous studies carried out in laboratories and real classrooms have provided similar results (i e., Logan et al., 1991; Hardison, 2003; Lee & Lyster, 2015). This finding further confirms the facilitating role of feedback in L2 speech learning in real classrooms. Nevertheless, the participants who were given feedback in the study conducted by Lee & Lyster (2015) seemed to have achieved higher post-test and delayed test accuracy. This might be explained by the fact that the participants had higher perception accuracy in the pre-test than those in the present study. Furthermore, with the exception of pronunciation-focused instruction, more forms of exercise were used by Lee & Lyster (2015) including pick-up-a-card games, bingo games and fill-in-the-blank exercises. The present study, however, only had one form of awareness task. To make up for this limitation, the participants in the present study went through a much longer time period of four weeks of instruction lessons. Given that *feedback* as a between-subjects factor was found to be significant for the participants' perception performance, we could speculate that the participants who were given feedback could have achieved higher accuracy if more awareness tasks were carried out in the instruction lessons.

Another interesting finding was that the participants' length of English study was revealed to have played non-significant effect on their perception performance in post-test and delayed test. If the amount of learning experience does play a critical role in L2 acquisition as predicted by SLM and PAM-L2, why did the participants who had longer history of English study did not outperform those who had been learning English for a shorter period of time? It might be explained by the English educational system of China, which emphasises on grammar and vocabulary rather than listening and spoken English. The majority of the provinces of China did not include oral

English as part of the college entrance examination. In some provinces, English listening was even excluded from the college entrance examination. Thus, the students may not have much chance to enhance their listening and spoken English.

Moreover, factors which were investigated with questionnaires, such as motivation in English study, whether traveled to or lived in English speaking countries, etc., were all revealed to be insignificant for the participants' perception performance. Detailed examination of the collected questionnaire indicated that there was not a big difference among the participants' answers. For instance, the majority of them reported to have been learning English for the need of examinations. None of them had any chance to use English on a daily basis, or had ever traveled or lived in English speaking countries. More importantly, none of them reported had been practicing oral English at their spare time. In future studies, it would be interesting to further investigate the issue among students who spent time in English study at their spare time with those who did not.

5. Conclusion

The present study investigated the role of instruction with corrective feedback in early and late L2 learners' acquisition of a non-native contrast in the aspect of perception. The findings have provided us with clear answers; The OA of L2 English learning was revealed to be a non-significant factor regarding the participants' perception accuracy in the post-test and delayed test, meaning that both early and late L2 learners may be able to acquire an L2 sound that they initially have difficulty with. Moreover, participants who were given feedback in the instruction lessons displayed significantly higher perception accuracy than those who were not given feedback. There is further evidence in support of the effectiveness of instructional input with feedback in L2 speech learning in real classrooms.

References

Best, C. T. (1995 b). Learning to perceive the sound pattern of English. In Rovee-Collier, C., and Lipsitt, L. (Ed.), *Advances in infancy research*. Hillsdale, NJ: Ablex. 217-304.

Best, C. T. (1995a). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Com- monalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34). Amsterdam, the Netherlands: Benjamins. http://dx.doi.org/10.1075/lllt.17.07bes

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In Nusbaum, H. C. (Ed.), *The development of speech perception: the transition from speech sounds to spoken words*. MIT Press, 167-224.

Boersma, P., & Weenink, D. J. M. (2015). Praat: doing phonetics by computer (Version 5.3.64). Amsterdam: Institute of Phonetic Sciences of the University of Amsterdam. [Computer program]. Retrieved from http://www.praat.org/

Bohn, O.S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*, 303–328. http://dx.doi.org/10.1017/S0142716400008912

Doughty, C., & Varela, E. (1998). Communicative focus on form. *Focus on form in classroom second language acquisition*, *1*, 114-138.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585. http://dx.doi.org/10.1017/s0272263104040021

Flege, J. E. (1981). The phonological basis of foreign accent: A hypotheses. *TESOL Quarterly, 15*(4), 443-455. http://dx.doi.org/10.2307/3586485

Flege, J. E. (1987). The production of "new" and "similar" phonemes in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47-65.

Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In Heubner, T. and Ferguson, C., (Eds.), *Crosscurrents in second language acquisition and linguistic theory*.

Philadelphia: John Benjamins, 249-284. http://dx.doi.org/10.1075/lald.2.15fle

Flege, J. E. (1992a). Speech learning in a second language. In Ferguson, C., Menn, L., and Stoel-Gammon, C. (Eds.), *Phonological development: Models, research, and application*. Timonium, MD: York Press, 565-604.

Flege, J. E. (1992b). The intelligibility of English vowels spoken by British and Dutch talkers. *Intelligibility in speech disorders: Theory, measurement, and management*, *1*, 157-232.

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

Flege, J. E. (1995b). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics*, *16*, 425-442.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In Schiller, N. O., and Meyer, A. S. (Eds.), *Phonetics and phonology in language comprehension and production, differences and similarities*, 319-355. http://dx.doi.org/10.1515/9783110895094.319

Flege, J. E., Bohn, O.S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*, 437–470. http://dx.doi.org/10.1006/jpho.1997.0052

Flege, J. E., Munro, M. J., & MacKay, I. R., (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, *16*(1), 1-26. http://dx.doi.org/10.1016/0167-6393(94)00044-B

Flege, J., (1988). The production and perception of speech sounds in a foreign languages. In: Winitz, H. (Ed.), *Human Communication and Its Disorders, A Review*. Norwood, N.J.: Ablex, 224-401.

Fullana, N., & Mora, J. C. (2008). Production and perception of voicing contrasts in English word-final obstruents: Assessing the effects of experience and starting age. In *New Sounds 2007: Proc 5th International Symposium on the Acquisition of Second Language Speech*, 207-221.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495-522. http://dx.doi.org/10.1017/S0142716403000250

Havranek, G. (1999). The effectiveness of corrective feedback: Preliminary resultso fan empirical study. *Acquisition et interaction en langue etrangere: Proceedings of the English ERUSLA Conference*, *2*, 189-206.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech communication*, *47*(3), 360-378. http://dx.doi.org/10.1016/j.specom.2005.04.007

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, *126*, 866. http://dx.doi.org/10.1121/1.3148196

Jamieson, D. G., & Rvachew, S. (1992). Remediating speech production errors with sound identification training. *Journal of Speech-Language Pathology and Audiology*, *16*(3), 201-210.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley and sons.

Li, Y., (2015). *Audio-visual Training Effect on L2 Perception and Production and Production of English /θ/-/s/ AND /ð/-/z/ By Mandarin Speakers*. Doctoral Dissertation. Newcastle University.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/. *The Journal of the Acoustical Society of America, 89*(2), 874-886. http://dx.doi.org/10.1121/1.1894649

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37-66. http://dx.doi.org/10.1017/s0272263197001034

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language and Hearing Research*, *40*(3), 686-693. http://dx.doi.org/10.1044/jslhr.4003.686

McGuire, G. (2010). A brief primer on experimental designs for speech perception research. *Laboratory Report*, *77*.

Ohta, A. (2000). Rethinking recasts: A learner-centered examnation of corrective feedback in the Japanese classroom. In J.K. Hall & L. Verplaetse (Eds.), *The construction of second and foreign language learning through classroom interaction* (pp.47-71). Mahwah, NJ: Erlbaum.

Oliver, R. (2000). Age differences I negotiation and feedback in classroom and pairwork. *Language* Learning, *50*,

119-151. http://dx.doi.org/10.1111/0023-8333.00113

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, *5*(3), 261-283. http://dx.doi.org/10.1007/BF01067377

Russell, J., & Spada, N. (2006). The Effectiveness of Corrective Feedback for the Acquisition of L2Grammar: A Meta-analysis of the Research. In J. Norris and L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Amsterdam: John Benjamins Publishing.

Shi, L. F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language and Hearing Research, 53*(4), 821. http://dx.doi.org/10.1044/1092-4388(2010/09-0081)

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989-994. http://dx.doi.org/10.1016/S0028-3932(02)00316-0

Williams, G.C., & McReynolds, L.C. (1975). The relationship between discrimination and articulation training in children with misarticulations. *Journal of Speech and Hearing Research*, *18*(3), 401-412. http://dx.doi.org/10.1044/jshr.1803.401

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

Notes

Note 1. Due to speaker differences, Mandarin /s/ can be pronounced as alveolar (Chang et al., 2009), apical or dental-alveolar (Lee, 1999).

Note 2. V=vowel; C=consonant

Note 3. *d-prime* scores were calculated with the formula *d'*=NORMINV(hit-rate, 0, 1) - NORMINV(false-alarm-rate, 0, 1). The highest possible *d'-prime score* (greatest sensitivity) is 6.93, and the effective limit (using .99 and .01) is 4.65 (see Li, 2015).

Appendix I

Questionnaire

1. Name (**姓名**):
2. Gender(**性**别) :
3. How old were you when you started English learning?(你几岁开始学英文的?)
4. How many years have you been learning English?(你学习英文多少年了?)
5. What is your primary motivation of learning English? (你学习英文的主要动机是什么?)
6. Have you ever travelled or lived in English speaking countries?(你曾经去过英语国家旅游或居住吗?)
7. Do you have any chance to use English on a daily life? (日常生活中你有机会使用英文吗?)
8. Except English classes at the university, do you study English at your spare time? If Yes, in which ways? (除了大学开设的英文课，你在课余时间学习英文吗？如果有，都有哪些学习形式?)

Appendix II

Nonsense words for perception tests

/θi/ /si/ /θa/ /sa/ /θu/ /su/ /iθi/ /isi/ /aθa/ /asa/ /usu/ /uθu/ /iθ/ /is/ /aθ/ /as/ /uθ/ /us/