Effects of recasts and prompts on L2 pronunciation development: Teaching English /ɹ/ to Korean adult EFL learners

Reginald Gooch [a, *], Kazuya Saito [b], Roy Lyster [a]

[a] Department of Integrated Studies in Education, McGill University, Canada
[b] Department of Applied Linguistics and Communication, Birbeck, University of London, UK

ARTICLE INFO

ABSTRACT

The current study aimed to test how two different types of corrective feedback—recasts and prompts—can differentially affect the pronunciation development of English /ɹ/ by Korean learners in the context of simulated meaning-oriented classrooms receiving form-focused instruction (FFI). Twenty-two Korean learners of English as a foreign language were divided into three groups (FFI-only, FFI-recasts, FFI-prompts). Each group received 4 h of instruction in which they were taught how to structure an argument in English through activities designed to simultaneously focus their attention on /ɹ/, including explicit articulatory instruction. All lessons were video-recorded, which allowed all instances of feedback and learner response to be identified and coded. Production was measured through both spontaneous and controlled tests, and judged by five trained English native-speaking listeners. According to the results of the pre/post-tests, recasts were especially helpful in the improvement of controlled production of /ɹ/, whereas prompts were facilitative of not only controlled but also spontaneous production of /ɹ/. The results of the video-coding analysis further demonstrated that, during instruction, students were pushed by prompts to improve intelligibility mainly through the adjustment of interlanguage strategies (e.g., prolonging the phonemic length), and by recasts to refine accuracy in their /ɹ/ production.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Corrective feedback (CF) plays a pivotal role in second language (L2) learning, contributing to better overall L2 performance than similar instruction without CF (Lyster, Saito, & Sato, 2013). Over the past 20 years, a great amount of research attention has been directed towards examining the differential effects of recasts (i.e., a reformulation of the student's errors) and prompts (i.e., a signal of the student errors, but without provision of correct forms) on the learning of L2 morphosyntax, as research of this kind entails much theoretical value (the role of positive and negative evidence) as well as pedagogical value (the manner/timing of drawing students' attention to form within a primarily communicative environment) in instructed second language acquisition (SLA). Given that such CF episodes have been reported to occur as a result of not only morphosyntax, but also pronunciation-related errors in various classroom settings, researchers have recently begun to expound

---

the potential of form-focused instruction (FFI) and CF for phonological targets. In our precursor research (Saito, 2013a; Saito & Lyster, 2012a), for example, we tested and confirmed the positive impact of one single type of phonological CF—recasts—on the development of various phonological targets by Japanese learners of English. This research has shown promise, but has yet to tease apart the effectiveness of various components of FFI and CF for phonological learning. Extending our previous research framework with Japanese learners of English, the current study investigates the generalizability of FFI and CF across different learning contexts—intermediate- to advanced-level Korean learners of English. At the same time, to our knowledge, the current study is one of the first attempts in the field of SLA that took an exploratory approach towards the disentanglement of CF effects for the phonological target English /ɹ/, focusing on recasts *versus* prompts.

2. Background

2.1. Pronunciation teaching

A growing number of studies have tried to find effective means to teach pronunciation in the classroom, as good phonological development is important for many L2 learners (Thomson & Derwing, 2015). Not only does it affect their ability to effectively communicate but has also been linked to orthographical development (Wang, Park, & Lee, 2006), feelings of belonging (Gluszek, 2010), confidence (Tang, Zhang, Li, & Zhao, 2013), willingness to communicate (Derwing, Munro, Foote, Waugh, & Fleming, 2014), and overall language development (French, 2006). The present study focuses on developing effective pedagogical techniques for pronunciation teaching in a Korean EFL context.

Korean learners of English often experience difficulty with English /ɹ/ (Cho & Park, 2006), and this difficulty can persist even after years of living in an English-speaking environment (Tench, 2003). Yet teachers have relatively little guidance for helping their students overcome this barrier. Indeed, researchers still make a strong call for the development of research-based pedagogy and teaching materials (Foote, Holtby, & Derwing, 2011). This is not to say that theories of L2 phonological learning do not exist, but that there is no clear path for teachers to follow when applying this theoretical knowledge to the real-world classroom. In the following sections, we will explore the literature concerning pronunciation instruction, CF, and Korean learners of English /ɹ/, ultimately laying out the rationale for this study.

A number of laboratory-based experiments have been able to successfully contribute to L2 learners' phonological development (for review, see Hardison, 2010). Despite gains shown in these laboratory studies, because of their decontextualized teaching techniques (e.g., hours of intensive exposure to minimal pairs without any contextualized use of language) and their generally isolated and artificial nature, they leave us with few teaching implications (Fraser, 2011). This is especially true as many of these studies (e.g., Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997) did not test whether learners were able to generalize their gains from a controlled to a spontaneous level. While it may be more difficult to design tests that accurately measure pronunciation in a more free and communicative context, this type of test has greater authenticity and ecological validity than controlled tests (Lee, Jang, & Plonsky, 2014).

2.2. FFI and CF

Over the past 25 years, a number of researchers have investigated how teachers can apply a more pedagogically-oriented approach—FFI—in the context of pronunciation teaching. FFI is defined as "any pedagogical effort which is used to draw the learners' attention to language form either implicitly or explicitly" (Spada, 1997, p. 73). Different from traditional teaching methods (e.g., audio-lingual, grammar translation methods) which introduce language forms in a decontextualized manner, FFI is considered to be most effective when it is implemented in content-based and meaning-oriented classrooms (Spada, 2011). As such, FFI is designed to help L2 learners use target language structures at not only a controlled but also a spontaneous speech level (Spada & Tomita, 2010), with a view to transferring what they have learned in the classroom to outside of the classroom (Trofimovich & Gatbonton, 2006). The effectiveness of FFI has been observed in many language domains including grammar (Spada & Tomita, 2010), vocabulary (Schmitt, 2008), pragmatics (Nguyen, Pham, & Pham, 2012), and pronunciation (Lee et al., 2014).

The instructional options in FFI have been operationalized in various ways. For example, explicit instruction on target structures is theorized to help orient students' attention to phonetic information, thereby giving students valuable declarative knowledge in order to accelerate the creation of new linguistic categories (DeKeyser, 2003). It is recommended to subsequently engage students in form-focused tasks, which are designed to guide L2 learners to use target phonetic structures while achieving the communicative purposes of those tasks (Ellis, 2006).

A crucial aspect of FFI in response to learner errors involves CF. Oral CF is defined as teachers' spoken responses to their students' linguistic errors. Much research attention has been given to examining which type of CF can impact SLA processes in an optimal fashion, especially focusing on recasts *versus* prompts (pushing students to self-repair without providing correct forms). The latter CF moves can be executed via metalinguistic clues, repetition, clarification requests, and elicitation. From a theoretical perspective, recasts can provide both positive evidence (provision of model forms) and negative evidence (a signal of error occurrences), whereas prompts entail only negative evidence (Lyster & Saito, 2010).

According to previous comparative studies of CF types, more gains have been identified for prompts than for recasts in classroom contexts of teacher-student interaction (ibid.). One reason is that the corrective purpose of recasts (i.e., negative evidence) can be ambiguous, especially when they are given in response to morphosyntactic errors, since those errors affect

communication less than vocabulary or pronunciation errors (Mackey, Gass, & McDonough, 2000). In contrast, prompts generally provide more salient negative evidence, which in turn elicits a greater amount of L2 learners' modified output (Ellis & Sheen, 2006).

2.3. Motivation for current study

Although much research evidence on FFI and CF has mainly concerned L2 morphosyntax development, some researchers have expanded its scope to other linguistic domains (for a comprehensive review, see Loewen, 2015). With respect to L2 pronunciation development (the focus of this paper), earlier descriptive studies have shown that L2 learners are sensitive to pronunciation-focused CF, resulting in a relatively large amount of immediate uptake and repair (Lyster, 1998; Sheen, 2006). More recent research has taken an exploratory approach towards examining the extent to which CF and FFI can help L2 learners with various first language (L1) backgrounds improve their phonological performance, such as Japanese speakers learning English vowels and consonants (Saito & Lyster, 2012a, 2012b), Korean speakers learning English vowels (Lee & Lyster, 2015), and Cantonese speakers learning Mandarin tones (Saito & Wu, 2014).

The results have generally shown that FFI can impact L2 learners' controlled and spontaneous pronunciation ability. The findings suggest that the efficacy of CF is related to learners' proficiency levels. On the one hand, adding CF to FFI lessons can be effective especially for learners with already-acquired phonetic knowledge (via explicit phonetic instruction or/and a few years of immersion in an L2 speaking environment). On the other hand, CF does not seem to make much difference for learners with limited L2 phonetic knowledge, arguably because these learners may be developmentally unready to process such production-based practice during FFI.

These studies have raised a range of questions worthy of future investigations.

One such issue involves the role of different types of CF. Notably, previous studies of the effects of CF provided on students' production errors (as opposed to perception; see Lee & Lyster, 2015) have examined only recasts in response to nontarget pronunciation (e.g., Saito & Lyster, 2012a, 2012b), rather than investigating how recasts and prompts *differentially* help L2 learners' pronunciation ability. In this regard, the current study was designed to compare three treatment groups—FFI-only, FFI-recast, and FFI-prompt—in the context of adult Korean learners' English /ɹ/ pronunciation development with a pre- and post-test design. With respect to the CF treatment, recasts were operationalized as partial recasting, and prompts as clarification requests and elicitations (for details, see the Method section). As such, we aimed not only to examine the role of CF in L2 phonetic acquisition but also to tease apart composite functions of CF effectiveness (positive vs. negative evidence). That is, students in the FFI-recast group received both negative and positive evidence, while those in the FFI-prompt group received only negative evidence.

Addressing this issue is not only theoretically relevant but also of pedagogical concern. For example, it is important for teachers to know which type of CF (recasts or prompts) to provide, and also when to provide different kinds of CF according to various stages of L2 learners' interlanguage development. For L2 acquisition of English /ɹ/, previous research has demonstrated that L2 learners initially pay attention to the relatively long length of the phoneme (>50 ms). They then become aware of the complex articulatory structure of the sound (three constrictions in labial, alveolar, and pharyngeal areas of vocal tract) (Bradlow, 2008). Importantly, whereas the former aspect of English /ɹ/ development (i.e., temporal change) is linked with native speakers' (Ns's) perceptual judgment of intelligibility (whether the sound can be considered as English /ɹ/ or other consonant sounds), the latter aspect of English /ɹ/ development (i.e., spectral change) directly impacts Ns's accuracy judgment (how accurately, easily, and promptly the sound can be perceived *within* the category of English /ɹ/) (Saito, 2013b).

From a pedagogical standpoint, it can be hypothesized that teachers may need to use different CF techniques—recasts or prompts—to help Korean learners acquire different aspects of English /ɹ/ production (prolonging the phonemic length → learning new articulatory parameters), depending on whether the learners' current interlanguage entails production of unintelligible /ɹ/, intelligible /ɹ/, or nativelike /ɹ/. Therefore, the following research question and predictions were formulated:

- How do different CF types differentially impact the different stages of L2 pronunciation development of /ɹ/?

With respect to the early phase of English /ɹ/ development (unintelligible to intelligible /ɹ/), it was predicted that Korean learners would benefit from prompts that provide negative evidence only, but that they may need the positive evidence available in recasts in order to attain nativelikeness in their /ɹ/ pronunciation (intelligible to nativelike /ɹ/).

3. Method

The present study followed a quasi-experimental pre- and posttest design in a simulated classroom setting at a public university in South Korea. Given that L2 learners acquire different domains of English /ɹ/ at different learning rates (Korean /ɾ/ → hybrid forms → intelligible English /ɹ/ → more targetlike English /ɹ/), our study attempted to investigate the complex relationship between recasts, prompts, and English /ɹ/ interlanguage development from various perspectives including online (i.e., in-class) performance and offline (i.e., pre/posttest) performance.

Participants were randomly assigned to either the FFI-only group, the FFI-recast group or the FFI-prompt group.[1] Each group followed the same procedure of pretest, treatment, and posttest. The treatment consisted of a 4-h language course, which was distributed over four separate 1-h sessions during a two-week period. Each class session in this study was video-recorded for later analysis of CF (the frequency of recasts and prompts) and student uptake following CF (successfully repaired, partially repaired, needs repair).

3.1. Student participants

The participants in this study were 22 undergraduate Korean learners of English as a foreign language (10 male, 12 female) recruited from a range of faculties at a public university in South Korea. Participants were recruited through flyers, which were posted on the bulletin board in the Humanities Building, and through an ad, which was posted on the English Department's webpage. Participants' ages ranged from 18 to 30 years. They were randomly assigned to one of three groups: the FFI-only group ($n = 9$), the FFI-recast group ($n = 7$), and the FFI-prompt group ($n = 6$). Initially, 25 students had been recruited, but 3 students withdrew from the study after groups had been assigned, resulting in slightly uneven distribution of participants.

3.2. English NS judges

Five English NS listeners (2 male, 3 female) were recruited from a university in Canada to rate 704 speech tokens recorded during pretests and posttests. All judges were NSs of English under the age of 50, and had at least two years of EFL or ESL teaching experience and normal hearing.

3.3. Instructor

The instructor for all classes was the first author (a male NS of American English with two years of experience as an EFL instructor at private language institutes in South Korea). In order to reduce instructor fatigue and ensure quality of instruction for all groups, classes were spaced out so that the instructor taught no more than two lessons per day. As an additional measure, the order in which the groups received instruction for each class day was alternated so that no one group would be taught consistently before or after the other groups.

3.4. Target of instruction

Korean L1 speakers tend to have difficulty with English /ɹ/. According to the Perceptual Assimilation Model (Best, 1995), Korean perception of English /ɹ/ and /l/ should fall into either the "single-category" (two L2 sounds are perceived as one L1 sound) or "category-goodness" (one of the two L2 sounds is more similar to the L1 sound) pattern, indicating that Koreans will face perceptual difficulties with these sounds. This theoretical prediction is echoed in empirical data. For Koreans, issues with pronunciation and perception of /ɹ/ are persistent, often continuing even after many years living in Anglophone countries (Tench, 2003). While there are many potential groups of learners and language targets which should be investigated, it is especially interesting to investigate Korean learners of /ɹ/ not only because of the well-documented difficulty this target presents, but also because Koreans constitute such a large population of EFL students.

3.5. Instructional treatment

The classroom instruction for all three groups consisted of FFI including explicit articulatory instruction. The classroom instruction also included CF for two of the three groups but not for the third.

3.5.1. Form-focused instruction

The instructional materials were adapted from our precursor research (Saito, 2013a; Saito & Lyster, 2012a) for use with Japanese learners of English. First, learners in all three groups received explicit instruction on English /ɹ/ for 10 min at the beginning of class on the first day, and a short review in the first 3–5 min of class on all other days. Explicit instruction included an introduction to the /ɹ/ sound, with an exaggerated model of pronunciation from the instructor (1 min), followed by an introduction to the different aspects of articulation of /ɹ/. For each articulatory aspect, the instructor explained what happens inside the mouth using cross-sectional charts, and provided exaggerated demonstrations.

Subsequently, students practiced English /ɹ/ in meaning-oriented activities. The primary content of the course was teaching English argumentation skills, but by giving the course through FFI, learner attention was focused on /ɹ/ in a number of ways. Over the course of four days, students learned to critique an argument, to support their points during a debate, to

---

[1] We did not include a control group that received neither FFI nor CF. Our precursor research (Saito, 2013a) had shown that such control groups demonstrate little improvement, the 4 h of intensive exposure to meaning-oriented instruction being insufficient to trigger perceptible changes in the pronunciation of English /ɹ/.

Table 1
39 Target words included in the instructional treatment.

Phonetic contexts	Target words
Word−initial	*Race, rain, ram, rat, read, reef, rent, right, rice, ride, rink, river, road, roan, robot, rock, rocket, Rome, roof, room, round, rule, run, Ryan, wrong, wrap*
Word−medial	*Arrive, correct, pirate*
Consonant cluster	*Bread, crab, crime, crowds, fries, fruit, grass, green, free, pray*

structure an argument, and ultimately, to create an argument and present it orally in front of the class. Throughout these activities, students encountered and used many target /ɹ/ words, which were typographically enhanced (bolded, red, and in italics) in the class. For example, students critiqued the argument "Whenever I eat Korean **rice** I have digestion problems. So, eating Korean **rice** causes digestion problems," and debated the topic "Is **read**ing comic books good for children or not?"

In addition to these main activities, students also played a variety of warm-up games designed to help them practice distinguishing English /ɹ/ in their perception and production. One example of the games used is English Karuta. In this game, 36 cards featuring pictures of minimally paired /ɹ/ or /l/ words were placed on the table. When the teacher said a word, students had to compete to grab the correct card as quickly as possible. At the end of the game, the student with the most cards was the winner. This game was designed to draw students' attention to the perceptual differences between /l/ and /ɹ/. A list of the target words appears in Table 1.

3.5.2. Corrective feedback

Two groups (FFI-recast and FFI-prompt) received different types of CF from the instructor.

3.5.2.1. FFI-recast group. While there are a number of different forms a recast can take, for the purpose of this study the instructor provided recasts in the following way: Whenever a student produced a non-targetlike /ɹ/ word, the instructor recast the word in a targetlike fashion with falling intonation. This combination of partial (one-word) recast and falling intonation has been shown to be more salient than other types of phonological recasts (Sheen, 2006), and even to predict posttest accuracy (Loewen & Philp, 2006). The following is an example of recast provision during one of the warm-up games (i.e., a guessing game) from Day 3. In this game, students formed pairs, and each pair was given a stack of cards. On each card was a target word. One student gave clues to his/her partner and attempted to get the partner to guess the word.

Student 1: Protects a house from rain and snow.
Student 2: /ruf/
Instructor: /ɹuf/ (with falling intonation)

3.5.2.2. FFI-prompt group. The most easily identifiable difference between prompts and recasts as phonological CF is that, while recasts provide positive evidence (i.e., the participants can hear the instructor's model pronunciation of the word), prompts do not. Like recasts, a wide variety of prompt types are possible. For the purpose of this study, the combination of a clarification request (e.g., "Pardon?") followed by an elicitation (e.g., "Could you say that again please?") was used. This combination will be referred to as a prompt throughout the rest of this study. The following is an example of prompt provision from Day 2, during a warm-up game of Karuta. After students heard their instructor reading aloud a target word, they tried to find a corresponding card. When they got the correct one, they were encouraged to pronounce the word which was on their card.

Instructor: Who can find 'pray?'
Student: /preɪ/
Instructor: Sorry? Can you say that again?
Student: /pɹeɪ/

3.6. Outcome measures

The tests used in this study were those used by Saito (2013a) and Saito and Lyster (2012a, 2012b) in their studies with Japanese learners of English. There was a pretest and a posttest, each of which had both a controlled production test and a spontaneous production test. Test items included both trained and untrained items in a variety of phonetic contexts. Trained items were defined as those words that appeared in the instructional materials and untrained as those that did not. Using a mixture of trained and untrained items allowed us to examine whether students were able to improve merely in the case of specific "r" words, or whether they were able to isolate their knowledge of the sound /ɹ/ and transfer it to other words. The participants took the spontaneous production test first, followed by the controlled production test.

In addition to the pre- and posttest measures, video recordings of both the recast and prompt classes were analyzed for the type and quantity of CF given and for the students' immediate responses (i.e., uptake).

3.6.1. Spontaneous production test

To measure spontaneous production, a picture description task was employed. This type of task is theorized to contribute more information about the linguistic target within global speech patterns than a controlled production task alone would do (Thomson & Derwing, 2015). The spontaneous production test measured student production of eight target r-initial words (four trained and four untrained), with a variety of vowel sounds following the initial /ɹ/. The initial /ɹ/ context was chosen for the purpose of comparability with previous studies investigating production of /ɹ/. The words used were as follows: *read*, *roof*, *rain*, *road* (trained), *ring*, *red*, *route*, *rope* (untrained). This spontaneous production test was designed to measure student production under time pressure with their primary focus on meaning rather than form. This kind of task can often be more cognitively challenging than production under controlled circumstances wherein the speaker does not have to attend to meaning.

The spontaneous production task was administered as follows. First, participants were given two sets of two words, some of which were target words and some of which were distracter items. Next, the participants were given a short time (<10 s) to memorize those words. Third, the words were taken away and the participants were given a set of two pictures (one at a time). Students then spontaneously described the picture using the given words. In total, the learners described 16 pictures using eight target words and various distracter words. All student responses were recorded by a Roland 05 Wave recorder.

3.6.2. Controlled production test

Like the spontaneous production test, the controlled production test featured eight target words. The words used were as follows: *rink*, *rule*, *race*, *road* (trained), *reach*, *rude*, *rate*, *roll* (untrained). The selected words feature word initial /ɹ/ followed by vowels of varying height and backness. For the task itself, participants were asked to read a list of 40 words, throughout which the eight target words were interspersed. In this way, students were able to read the words without having to attend to morphosyntactic, semantic or other concerns. Responses were recorded with the Roland 05 recorder.

3.6.3. Native speakers' judgment of production tests

During the spontaneous and controlled production tests a total of 704 items were recorded (22 students × 2 test sessions × 16 items = 704). These items were extracted and isolated using the computer software *Praat* which allows researchers to carefully check where each phoneme starts and ends based on spectrogram images as well as waveforms. This was necessary especially in the case of the items from the spontaneous production test, which were produced with their primary focus on meaning rather than form, and target words were embedded in a continuous speech stream. In order to isolate these words, the first author listened multiple times to find the onset of each target word (i.e., a sharp decline in the third formant). Any inflected endings for target words (e.g., roads, raining) were included to avoid distortion.

After all 704 items had been prepared in this way, the items were randomized for judgment by NS judges. Judgments were completed individually in a quiet room at a Canadian university. During judgment, tokens were presented one-by-one to the NS judge on a computer screen as he/she listened to a sound file through headphones. Tokens were judged on the same 9-point scale originally elaborated in Flege, Takagi, and Mann (1995), that is, *1 (Very good /ɹ/)*, *2 (Good /ɹ/)*, *3 (Probably /ɹ/)*, *4 (Possibly /ɹ/)*, *5 (Neutral)*, *6 (Possibly /l/)*, *7 (Probably /l/)*, *8 (Good /l/)*, *9 (Very good /l/)*. To make a judgment along this scale, the judge simply clicked the corresponding button onscreen. There was also a repeat button available, which could be clicked a maximum of three times per token. While making judgments, NS judges were asked as much as possible to base their judgments on the initial sound for each word they heard, rather than to take other factors into account, such as the mispronunciations of other sounds within the words. To judge the entire dataset, each judge took between 60 and 90 min.

3.7. Video-coding analysis

The video recordings of both the FFI-recast and FFI-prompt groups were analyzed to track instances of CF and student uptake in response to the CF. Lyster and Ranta (1997) defined uptake as "a student's utterance that immediately follows the teacher's feedback and that constitutes a reaction in some way to the teacher's intention to draw attention to some aspect of the student's initial utterance" (p. 49). They classified uptake as learner utterances with repair and those still in need of repair. We used a modified version of their coding scheme as follows:

- Student does not attempt repair (i.e., no uptake).
- Student attempts repair but responds with Korean /r/.
- Student attempts repair but produces a hybrid form with elements of English /ɹ/ and Korean /ɾ/.
- Student repairs using English /ɹ/.

To account for interlanguage development, this coding scheme includes student responses that could be classified as neither /ɹ/ nor /r/. Although this classification accounts for a variety of student responses, we do not necessarily take the stance

Table 2
Descriptive results of the production tests.

Group		Controlled test				Spontaneous test			
		Trained lexical items		Untrained lexical items		Trained lexical items		Untrained lexical items	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post
FFI-Recasts	M	2.17	2.31	2.31	1.89	2.71	2.37	3.17	2.80
($n = 7$)	SD	0.66	0.76	0.75	0.72	0.96	1.08	0.73	0.84
FFI-Prompts	M	3.30	2.52	3.48	2.52	3.68	3.00	3.98	2.92
($n = 6$)	SD	1.18	0.62	1.53	0.64	1.21	0.57	1.37	0.76
FFI-only	M	2.44	2.33	2.82	2.57	3.10	2.90	3.56	3.38
($n = 9$)	SD	0.62	0.59	0.86	0.60	0.74	0.85	1.12	0.83

that all responses in this range represent equal progress in students' L2 phonological development. We used this classification in order to be able to note the instances in which students attempted to repair from /ɾ/, but were unable to produce targetlike /ɹ/.

To ensure the accuracy of data coding, the first author and another coder (an experienced L2 instructor) independently coded all classroom data for both the recast and prompt groups. After conducting independent analyses, the researcher and coder discussed every instance of CF and attempted repair until they were able to reach a consensus on the coding.

4. Results

Results are divided into two main sections. The first section will detail the overall results from the pretests and posttests, and present a short statistical analysis. The second section will look at instances of CF and/or learner repair for the two experimental groups, and analyze trends in those groups.

4.1. Pre- and posttest results

In this subsection, we examine whether and to what degree FFI with and without CF had an impact on learners' phonological improvement in production over time. Raw production test scores are summarized in Table 2, revealing that most of the participants already demonstrated relatively advanced phonological skills at the time of the pre-tests in controlled ($M = 2.17$–3.48) and spontaneous production ($M = 2.71$–3.91) (recall that the range from highest to lowest possible score is 1–9).

Because the visual inspection of histograms revealed positive skewness, the data were transformed via log transformation to approximate log-normal distributions. To ensure the validity of any inferential statistics based on the small sample size ($n = 22$), we also report statistical power for each significant effect.

The production test scores were submitted to a 4-way ANOVA with Group (FFI-only, FFI-recast, FFI-prompt) as a between-subject factor and Task (controlled, spontaneous), Lexis (trained, untrained items), and Time (pretests, posttests) as within-subject factors. Although no significant main or interaction effects were found for Group or Lexis ($p > 0.05$), there were significant main effects for Task, $F(1, 19) = 36.577$, $p < 0.001$, with large statistical power (1.00), and for Time, $F(1, 19) = 8.916$, $p = 0.008$, with large statistical power (0.81). According to Bonferroni multiple comparisons, whereas the recast group demonstrated significant improvement only in the context of untrained lexical items in the controlled task ($p = 0.046$), the prompt group demonstrated similar generalizable gains both in the controlled ($p = 0.017$) and spontaneous tasks ($p = 0.036$).[2]

Taken together, these results are suggestive of the following patterns, which remain tentative due to the small sample size. First, the FFI-only treatment (without CF) was not sufficient to trigger a significant positive change in the participants' /ɹ/ pronunciation. Second, the recast group's gain was evident in their controlled /ɹ/ production of untrained lexical items. Third, the prompt group's improvement was apparent not only in the case of controlled production of untrained lexical items but also in more spontaneous levels of speech processing.

4.2. Corrective feedback and learner responses

All instances of CF and student uptake were tracked and coded throughout this study according to the four categories (no uptake, Korean /ɾ/ repair, hybrid repair, English /ɹ/ repair). In this subsection, we present the results of this coding for the two CF groups.

---

[2] In the current study, all Bonferroni-adjusted significance tests were conducted via SPSS by multiplying an unadjusted p value by the number of comparisons.

Table 3
Descriptive data of CF and learner uptake in the FFI-recast group.

	Day 1	Day 2	Day 3	Day 4	Total
Instances of CF	17	3	40	8	68
Instances of uptake	11	3	33	7	54
Uptake/CF ratio	65%	100%	83%	88%	79%

4.2.1. FFI-recast group

As displayed in Table 3, students in the FFI-recast group received a total of 68 recasts and produced 54 instances of uptake, indicating repair or attempts at repair following 79% of the recasts. As seen in Table 4, the student uptake following recasts included 6 responses (11%) with Korean /ɾ/, 15 responses (28%) with hybrid forms, and 33 responses (61%) with English /ɹ/. An analysis of student responses over time revealed that, even at the beginning of the study, this group rarely responded to recasts with Korean /ɾ/ (11%) and was often able to produce English /ɹ/ (82%), confirming that these participants could be considered relatively advanced learners. One additional noteworthy trend is that students attempted repair at a comparably much higher rate after Day 1 (see Table 3), suggesting that over the course of the instructional treatment, phonological recasts may have become increasingly salient to the learners.[3]

4.2.2. FFI-prompt group

As displayed in Table 5, students in the FFI-prompt group were provided a total of 97 instances of CF and they produced 90 instances of uptake, thus yielding an uptake rate of 93%. As seen in Table 6, of those 90 instances of uptake, 9 included Korean /ɾ/ (10%). Students produced hybrid forms 41 times (46%) and English /ɹ/ 40 times (44%).

It is interesting to note how learner responses in the FFI-prompt group changed over the course of the four-day program (Table 6). The most noteworthy trend that emerges is how dramatically attempts at repair resulting in Korean /ɾ/ drop off after Day 1. While uptake with /ɾ/ accounted for 26% of all student uptake on Day 1, this number dropped to 8% for Days 2 and 3, and zero on Day 4 of the study. Correspondingly, the number of repairs containing English /ɹ/ increased after Day 1. While 26% of responses on Day 1 included English /ɹ/, this rate was no lower than 33% on any of the following days. The pattern for uptake with hybrid forms is less clear, as the percentage of hybrid forms decreased on Days 2 and 3, but increased to their highest level on Day 4.

5. Discussion

By interfacing L2 speech and education research perspectives, the current study took an exploratory approach towards examining the role of FFI and CF in Korean leaners' English /ɹ/ pronunciation development. Our study is a novel contribution in that we tested the differential effectiveness of two types of CF (recasts, prompts) on their interlanguage development processes during the treatment (via video-coding analyses) and ultimate acquisition (via pre-posttest data analyses). First of all, this study has demonstrated that exposing L2 learners only to phonological FFI (consisting of explicit instruction and focused tasks) may not suffice to significantly impact their production of /ɹ/ between pretest and posttest. This is in line with our previous research which has shown that, whereas the FFI-only approach is beneficial for inexperienced L2 learners without much phonological knowledge of target sounds, adding output enhancement via provision of CF might be necessary especially for relatively advanced L2 learners, such as the Korean learners in the current study, to attain more nativelike pronunciation (Saito & Lyster, 2012a, 2012b). In the following sections, we address whether, to what degree, and how adding two types of CF—recasts and prompts—differentially affects the pedagogical potential of FFI at various stages of development of Korean learners as they move from Korean /ɾ/ to hybrid forms (with elements of both Korean /ɾ/ and English /ɹ/) to more targetlike production of English /ɹ/.

Overall, results of the statistical analyses showed that gains resulting from CF were particularly evident when participants' /ɹ/ performance entailed untrained rather than trained lexical contexts. Therefore, CF may have led the participants to establish, reinforce, and generalize their new phonological knowledge of English /ɹ/ beyond the lexical items that they had practiced during the treatment. Following recent adult L2 speech learning models (e.g., Bundgaard-Nielsen, Best, & Tyler, 2011), the results here suggest that CF is facilitative of the participants' attentional transition from vocabulary to sound learning. Another explanation for the complex results could be related to the fact that the participants may have had slightly more room for improvement in the former than the latter lexical contexts. According to the descriptive statistics in Table 2, it appears that the participants' /ɹ/ performance on the untrained items was less positively evaluated by the judges than their performance on the trained items.

Furthermore, the results also hinted at a continuum of effectiveness among the different CF conditions. To be precise, whereas the gains made by the recast group were observed especially at the controlled/untrained level of production, those of the prompt group seemed to be strong at both the controlled/untrained and the spontaneous/untrained levels of production.

---

[3] On Day 2, the instructor provided only three recasts, which indicates that the participants in the recast group did not make many pronunciation errors and thus could be considered relatively advanced.

Table 4
Amount (and percentage) of different types of uptake produced by the FFI-recast group.

	Day 1	Day 2	Day 3	Day 4	Total
Korean /ɾ/	1 (11%)	1 (33%)	4 (12%)	0 (0%)	6 (11%)
Hybrid form	1 (11%)	1 (33%)	11 (33%)	2 (29%)	15 (28%)
English /ɹ/	9 (82%)	1 (33%)	18 (55%)	5 (71%)	33 (61%)

Results of the video-coding analysis provide evidence that both phonological recasts and prompts can be salient for English L2 learners. Recasts used in this study led to student uptake following 79% of the recasts, which is consistent with the results of previous studies of learner responses following phonological recasts (Sheen, 2006). Even more salient were prompts, leading to uptake after 93% of all prompts.

It is important to mention here that the way students repaired (or attempted to repair) their non-targetlike English /ɹ/ differed according to CF type. The prompt group produced significantly more hybrid forms after CF than the recast group. Unlike recasts, which give students the option to mimic the instructor's pronunciation, prompts push students to use their own resources to try to produce a targetlike utterance (i.e., unintelligible to intelligible /ɹ/). That students from the prompt group produced a high percentage of hybrid forms containing elements of both Korean /ɾ/ and English /ɹ/, while subsequently showing improvement on posttests, suggests that those opportunities for modified output ultimately led to their more intelligible production of /ɹ/.

It remains to be seen what the limits of such a restructuring process might be. For example, learners in this study were able to draw on information provided during the FFI as a basis for trying to improve their interlanguage strategies. If students had not received explicit instruction, results from the prompt group may well have been less positive. In our previous research (Saito, 2013a), we indeed found that pronunciation-focused CF (recasts) can be effective especially when it is combined with explicit phonetic instruction prior to the FFI treatment.

Despite the prompt group's strong improvements in production, one potential weakness of prompts as a type of phonological feedback was also made apparent. After Day 1 of the study, the prompt group saw a dramatic decrease in its production of Korean /ɾ/ (26%—8%) and a concurrent rise in its rate of repair to /ɹ/ (26%—58%), but these levels plateaued after Day 2. This begs the question as to whether provision of positive evidence in CF might be needed at some point to help students reach higher levels of attainment.

Once students' development towards an English /ɹ/ stagnates due to over-dependence on interlanguage strategies, it might be beneficial for teachers to model the pronunciation (that is, to provide a recast) as a way to resolve this knowledge gap (Loewen, 2002). As suggested by researchers (e.g., Mackey & Philp, 1998), a recast can benefit students with partially-acquired L2 knowledge by providing not only negative but also positive evidence in order to further improve the accuracy of their interlanguage production. In the context of the current study, therefore, recasts may be effective for the later stage of L2 English /ɹ/ development (filling in the nativelikeness of already-intelligible /ɹ/ pronunciation).

In short, this study has demonstrated that prompts hold promise for helping learners improve their pronunciation of English /ɹ/ (unintelligible to intelligible /ɹ/), while also revealing some of the weaknesses of both prompts (insufficient for attaining nativelike /ɹ/) and recasts (insufficient for a more spontaneous processing ability). A mixture of CF types, rather than only one or the other, may ultimately prove to be the most effective for classroom application. Based on the results of this study, a promising combination would be the initial use of prompts to push learners to draw on their own resources (especially drawing on what they have learned through explicit instruction). Prompts can also push students to attend to target phonological forms even during meaning-oriented communication more explicitly than recasts (cf. uptake ratio = 93% for prompts and, 79% recasts in the current study). After students' noticing and awareness of target sounds is sufficiently enhanced, teachers can resort to other types of CF such as recasts especially when students are unable to self-correct or need additional help to move beyond interlanguage strategies to more targetlike production. In this way, the strengths of each type of CF can be combined while circumventing each one's weaknesses. Moreover, providing more than only type of CF is considered more feasible and realistic for teachers, as suggested by Lyster et al.'s (2013) earlier call for research on "combinations of CF types that more closely resemble teachers' practices in classroom settings" (p. 30).

6. Conclusion and future directions

Although there is still much work to be done in establishing best practices for pronunciation instruction, this study has taken a step towards establishing what those best practices might be. This study built on previous investigations into the efficacy of FFI for pronunciation teaching in the context of Japanese EFL, and provided evidence that those practices can be effective in the context of Korean EFL. Whereas previous studies focused on beginner-level students, this study demonstrated the feasibility of FFI with CF for improving L2 production even among advanced learners striving to achieve more accurate pronunciation. Additionally, this study points towards new possibilities in pronunciation-centered FFI by employing different types of CF (i.e., recasts and prompts) at different stages of L2 production development of /ɹ/ (unintelligible → intelligible → nativelike /ɹ/). Provision of CF in ways that research has shown to be effective is apt to have a positive impact on both teaching and learning. Specifically, our results suggest that, whereas both recasts and prompts have similar effects on L2 pronunciation development, prompts may result in quicker and more tangible improvement compared to recasts. At the same time, it is

Table 5
Descriptive data of CF and learner uptake in the FFI-prompt group.

	Day 1	Day 2	Day 3	Day 4	Total
Instance of CF	20	28	27	22	97
Instances of uptake	19	26	24	21	90
Uptake/CF ratio	95%	93%	89%	96%	93%

Table 6
Amount (and percentage) of different types of uptake produced by the FFI-prompt group.

	Day 1	Day 2	Day 3	Day 4	Total
Korean /r/	5 (26%)	2 (8%)	2 (8%)	0 (0%)	9 (10%)
Hybrid form	9 (47%)	9 (35%)	9 (38%)	14 (67%)	41 (46%)
English /ɹ/	5 (26%)	15 (58%)	13 (54%)	7 (33%)	40 (44%)

important to remember that the effectiveness of prompts may rest with helping learners to improve interlanguage strategies, but not necessarily to move beyond those interlanguage strategies. Learners may require additional help to do so, possibly in the form of CF with positive evidence.

As the current study was the first to investigate the effects of different types of CF in pronunciation teaching, our results provide a number of interesting questions for future research to address. For example, not only should the effect of different types of CF on learner acquisition of phonological targets be investigated, but also how those different types of CF can be best used, and in what combinations. One promising future direction involves the pedagogical potential of a combination of CF types that target errors in perception rather than production (Lee & Lyster, 2015). In addition, future studies should continue to investigate a diverse array of consonants and vowels, segmental and suprasegmental targets, not only for ESL/EFL, but also for a variety of target languages. Only when techniques for pronunciation instruction are investigated in a variety of environments can we begin to confirm the elements of instruction that are beneficial in a specific context and those that may apply across instructional settings.

For any future instructed L2 speech studies of this kind, we would like to emphasize the importance of adopting a range of outcome measures investigating the role of CF in L2 pronunciation learning from multiple angles. As L2 learners improve their interlanguage phonology by using different processing strategies (Major, 2008), the instruction-proficiency links need to be carefully scrutinized based on perception and controlled/spontaneous production tasks (Lee et al., 2015) according to various evaluation methods including human judgments (measuring the intelligibility and goodness of English /ɹ/) and acoustic analyses (measuring any change in spectral and temporal properties of English /ɹ/) (Lee & Lyster, 2015).

Finally, whereas the small sample size in the present study weakened the statistical design and limited the generalizability of results, it allowed us the opportunity to closely track individual learners during online production in class. In this respect, we have illustrated a useful way to examine not only the product of the CF effectiveness via pre/posttest measures, but also the process of instructed L2 speech learning via in-class assessments.

Acknowledgements

References

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

Bradlow, A. (2008). Training non-native language sound patterns. In J. Hansen, & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 287—308). Philadelphia, PA: John Benjamins.

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

Bundgaard-Nielsen, R., Best, C., & Tyler, M. (2011). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition, 33*, 433—461. http://dx.doi.org/10.1017/S0272263111000040.

Cho, J., & Park, H. (2006). A comparative analysis of Korean-English phonological structures and processes for pronunciation pedagogy in interpretation training. *Meta: Translators' Journal, 51*(2), 229—246. http://dx.doi.org/10.7202/013253ar.

DeKeyser, R. (2003). Implicit and explicit learning. In M. Long, & C. Doughty (Eds.), *Handbook of second language acquisition* (pp. 313—348). Malden, MA: Blackwell.

Derwing, T., Munro, M., Foote, J., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning, 64*, 526—548. http://dx.doi.org/10.1111/lang.12053.

Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review, 19*, 18—41. http://dx.doi.org/10.1075/aila.19.04ell.

Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition, 28*(04), 575—600. http://dx.doi.org/10.1017/S027226310606027X.

Flege, J., Takagi, N., & Mann, V. (1995). Japanese Adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38*(1), 25—55. http://dx.doi.org/10.1177/002383099503800102.

Foote, J., Holtby, A., & Derwing, T. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*, 1—22.

Fraser, H. (2011). Teaching teachers to teach /ɹ/ and /l/ to Japanese learners of English: An integrated approach. In M. Ashby (Ed.), *Proceedings of the Phonetics Teaching and Learning Conference* (pp. 11–15). Retrieved from https://www.ucl.ac.uk/psychlangsci/ptlc/ptlc2011-proceedings.

French, L. (2006). *Phonological working memory and second language acquisition: A developmental study of francophone children learning English in Quebec.* Lewiston, NY: The Edwin Mellen Press.

Gluszek, A. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology, 29*(2), 224–234. http://dx.doi.org/10.1077/0261927X09359590.

Hardison, D. (2010). Visual and auditory input in second-language speech processing. *Language Teaching, 43*, 84–95. http://dx.doi.org/10.1017/S0261444809990176.

Lee, J., Jang, J., & Plonsky, L. (2014). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics.* http://dx.doi.org/10.1093/applin/amu040. Advanced online publication.

Lee, A., & Lyster, R. (2015). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition, 38*. doi http://dx.doi.org/10.1017/S0272263115000194.

Loewen, S. (2002). *The occurrence and effectiveness of incidental focus on form in meaning-focused ESL lessons* (Doctoral dissertation). Retrieved from ResearchSpace@Auckland.

Loewen, S. (2015). *Introduction to instructed second language acquisition.* New York, NY: Routledge.

Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal, 90*(4), 536–556. http://dx.doi.org/10.1111/j.1540-4781.2006.00465.x.

Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning, 48*(2), 183–218.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition, 19*, 37–66.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition, 32*(2), 265–302.

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

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*, 471–497. http://dx.doi.org/10.1017/S0272263100004022.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses and red herrings? *Modern Language Journal, 82*, 338–356.

Major, R. (2008). Transfer in second language phonology: A review. In J. Hansen Edwards, & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 63–94). Amsterdam: John Benjamins.

Nguyen, T., Pham, T., & Pham, M. (2012). The relative effects of explicit and implicit form focused instruction on the development of L2 pragmatic competence. *Journal of Pragmatics, 44*, 416–434. http://dx.doi.org/10.1016/j.pragma.2012.01.003.

Saito, K. (2013a). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /ɹ/ to adult Japanese learners. *Language Learning, 63*(3), 499–529. http://dx.doi.org/10.1111/lang.12015.

Saito, K. (2013b). Re-examining effects of form-focused instruction on L2 pronunciation development: The role of explicit phonetic information. *Studies in Second Language Acquisition, 36*, 647–680.

Saito, K., & Lyster, R. (2012a). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning, 62*(2), 595–633. http://dx.doi.org/10.1111/j.1467-9922.2011.00639.x.

Saito, K., & Lyster, R. (2012b). Investigating pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly, 46*, 387–398.

Saito, K., & Wu, X. (2014). Communicative focus on form and L2 suprasegmental learning: Teaching Cantonese learners to perceive Mandarin tones. *Studies in Second Language Acquisition, 36*, 647–680.

Schmitt, N. (2008). State of the art: Instructed second language vocabulary acquisition. *Language Teaching Research, 12*, 329–363. http://dx.doi.org/10.1007/978-3-642-80027-6_2.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research, 10*, 361–392. http://dx.doi.org/10.1191/1362168806lr203oa.

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 29*, 73–87. http://dx.doi.org/10.1017/S0261444800012799.

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching, 44*, 225–236. http://dx.doi.org/10.1017/S0261444810000224.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*, 263–308. http://dx.doi.org/10.1111/j.1467-9922.2010.00562.x.

Tang, X., Zhang, S., Li, Y., & Zhao, M. (2013). Study on correlation of English pronunciation self-concept to English learning. *English Language Teaching, 6*(4), 74–79. http://dx.doi.org/10.5539/elt.v6n4p74.

Tench, P. (2003). Non-native speakers misperceptions of English vowels and consonants: Evidence from Korean adults in the UK. *International Review of Applied Linguistics in Language Teaching, 41*(2), 145–173. http://dx.doi.org/10.1515/iral.2003.006.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*, 326–344. http://dx.doi.org/10.1093/applin/amu076.

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in L2 Spanish word processing: Implications for pronunciation instruction. *The Modern Language Journal, 90*, 519–535. http://dx.doi.org/10.1111/j.1540-4781.2006.00464.x.

Wang, M., Park, Y., & Lee, K. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Educational Psychology, 98*(1), 148–158. http://dx.doi.org/10.1037/0022-0663.98.1.148.