Title:

Communicative Focus on L2 Phonetic Form: Teaching Japanese Learners to Perceive and Produce English /ɹ/ Without Explicit Instruction

Running Head:

INSTRUCTED L2 SPEECH LEARNING

Author Info:

Kazuya Saito

Waseda University

School of Commerce

1-6-1 Nishi Waseda,

Shinjuku, Tokyo

169-8050, JAPAN

Email: kazuya.saito@waseda.jp

Abstract

The current study examines in depth how two types of form-focused instruction (FFI)—FFI with and without corrective feedback (CF)—can facilitate L2 speech perception and production of /ɹ/ by 49 Japanese learners in EFL settings. FFI effectiveness was assessed via three outcome measures (perception, controlled production, spontaneous production) and also according to two lexical contexts (trained, untrained items). Two experimental groups received four hours of FFI treatment to notice and practice the target feature of /ɹ/ (but without any explicit instruction) in meaningful discourse. A control group (*n* = 14) received comparable instruction in the absence of FFI. During FFI, the instructors provided CF only to students in the FFI+CF group (*n* = 18) by recasting their mispronunciations of /ɹ/, while no CF was provided to those in the FFI-only group (*n* = 17). Analyses of pre- and post-tests showed that (a) FFI itself can sufficiently promote the development of speech perception and production of /ɹ/, and (b) the acquisitional value of CF in L2 speech learning remains unclear. The results suggest that the beginner learners without much phonetic knowledge on how to repair their mispronunciation of /ɹ/ should be encouraged to learn the target sound only through FFI in a receptive mode without much pressure for modified output.

Key words: Form-focused instruction, Second language phonetics, Pronunciation teaching, Listening teaching, English /ɹ/

Acknowledgement

This study was funded by the Government of Canada Post-Doctoral Research Fellowship and Waseda University Grant for Special Research Projects. I gratefully acknowledge insightful comments from Murray Munro, Roy Lyster, Tracey Derwing, Helen Fraser, and three anonymous *Applied Psycholinguistics* reviewers on earlier versions of the manuscript. Finally, I would like to extend my gratitude and appreciation to the administration staff, teachers, and students at ECC Foreign Language Institute, especially Noriaki Okamura, Yukio Nagato, and Anna-Marie Staniewski.

Over the past 25 years, second language acquisition (SLA) studies have extensively examined the effects of integrating form-focused instruction (FFI) into meaning-oriented classrooms as “a set of psycholinguistically motivated pedagogic options” (Ellis, 2001, p. 12) in the contexts of grammar teaching. They found that FFI can positively impact learners’ developing system of second language (L2) morphosyntax not only at a controlled level, but also at a spontaneous level (Norris & Ortega, 2000; Spada & Tomita, 2010). Although it is often assumed that the empirical findings of L2 grammar studies may be applicable to all types of language features such as L2 lexis (e.g., Schmitt, 2008) and L2 pragmatics (e.g., Rose & Kasper, 2002), it is surprising that little attention has been given to FFI research in the domain of L2 phonetic development. Building on the FFI research framework in the field of instructed SLA, the current study examines in depth which combination of FFI techniques –FFI with and without corrective feedback (CF)—can most benefit the development of phonetic aspects of L2 speech perception and production. We provided a range of pedagogical activities designed to elicit learners’ noticing and practice of the target sounds (but without any explicit instruction) in simulated classroom settings.

L2 Speech Learning

Theoretical Issues

Recently, speech researchers have proposed that L2 speech learning is subject to patterns of sound and word recognition that are similar to those of L1 phonetic acquisition (Best & Tyler, 2007; Bundgaard-Nielsen, Best, & Tyler, 2011a, 2011b; Walley, 2007). That is, L2 learners initially use computational strategies to detect prosodic patterns of language and then start recognizing words in the earliest stages of their vocabulary learning (i.e., lexically-driven L2 speech learning). Their representation of sounds, therefore, is based on lexical items that they have fully or partially acquired, subject to the influence of various lexical factors such as text frequency, subjective familiarity, and neighbourhood density (Bradlow & Pisoni, 1997; Trofimovich et al., 2012; Imai, Walley, & Flege, 2005).

As their vocabulary size increases, however, they are forced to engage in more fine-grained phoneme discrimination and identification in order to accurately comprehend and produce a large lexicon containing many phonetically similar words such as minimal pairs. Ultimately, while they are sensitive primarily to word-sized units of L2 speech information, they concurrently become more capable of detecting new sounds in L2 input at a phonetic level (i.e., phonetically-driven L2 speech learning). This phonetic-level restructuring, in particular, leads learners to create new phonetic categories and to generalize newly-acquired phonetic knowledge from familiar to unfamiliar lexical contexts (Flege et al., 1998). Finally, major L2 speech theories assume that phonetic categories have a perceptual basis and L2 learners access them at various processing levels. This suggests that (a) perception of new L2 sounds activates relevant sensorimotor skills and leads to production ability (i.e., a perception-first view: Flege, 1995, 2003; Kuhl, 2000) and (b) change appears first at a controlled and later at a spontaneous level (i.e., a variationist view: Major, 2008; Rau, Chang, & Tarone, 2009).

English /ɹ/

One well-researched yet still controversial topic in adult L2 speech learning is the acquisition of word-initial /ɹ/ by native speakers of Japanese (for a review, see Bradlow, 2008). From an acoustic point of view, American English /ɹ/ can be characterized along multiple dimensions, such as (a) third formant (F3), (b) second formant (F2), (c) first formant (F1) and (d) transitional duration of F1 and F3. A clear delimitation of /ɹ/ from the other approximant sounds in the English phonetic system requires, in particular, a severe dip in F3 as its reliable acoustic correlate (F3 = 1300-1950 Hz: Espy-Wilson et al., 2000). Despite a variety of articulatory configurations for the sound (e.g., retroflexed and bunched /ɹ/), key articulatory features include constrictions in the labial, palatal, and pharyngeal regions of the vocal tract and a front cavity resonance behind the sublingual space, which is highly related to low F3 values (Espy-Wilson et al., 2000).

The Japanese phonetic system includes neither /ɹ/ nor /l/, and Japanese learners tend to perceptually substitute the nearest L1 counterpart (Japanese tap /ɾ/) for both of these phonemes, thus neutralizing the contrast (Guion et al., 2000). Japanese learners resort to two interlanguage strategies instead of developing the new phonetic representation of F3 and relevant articulatory configurations. First, Japanese learners tend to overly depend on F2 frequencies to discriminate /ɹ/ from /l/ (Iverson et al., 2003) and produce /ɹ/ with lower F2 (tongue retraction, /w/-like production) (Saito & Brajot, 2013). Second, Japanese learners also overly rely on durational cues rather than spectral cues (Iverson et al., 2005). Although the transition duration of English /ɹ/ (50-100ms) is generally longer than that of Japanese tap /ɾ/ (5-20 ms), it has a wide range of natural variation (e.g., some /ɹ/ tokens can be as short as Japanese tap) and it is not necessarily a significant acoustic correlate of /ɹ/ (Flege et al., 1995b).

In sum, if Japanese learners are overly reliant on F2 distributions and temporal dimensions at the expense of variation in F3 (Saito & Brajot, 2013), the following framework for acquiring /ɹ/ may be proposed: (a) a move away from (though not necessarily a complete abandonment of) "default strategy" of F2 and duration dependency; (b) attention to new acoustic information (F3 variance: 2400-3000 Hz for Japanese tap /ɾ/ → 1600-1900 Hz for English /ɹ/); and (c) associated shift in orolingual articulation (i.e., labial, palatal and pharyngeal constriction).

In conjunction with the interlanguage development of /ɹ/, this study examines in depth whether and to what degree FFI can help L2 learners to expedite the development and restructuring of the multiple acoustic properties of their representations of /ɹ/ (i.e., F3, F2, F1, transition duration) as well as enhance various levels of their processing abilities of /ɹ/ (i.e., perception → controlled production → spontaneous production).

Pedagogical Issues

From a pedagogical perspective, developing effective and efficient approaches to helping adult L2 learners reach comprehensible pronunciation and listening skills is an especially timely and important initiative. Although foreign accents are normal aspects of L2 speech production for adult L2 learners (Flege, Munro, & MacKay, 1995a) and accents are essential to maintaining English users' sense of identity in the context of global communication (Jenkins, 2000; cf. Derwing & Munro, 2005), heavy accents may in some cases prevent even those with precise grammar from attaining successful comprehensibility in their L2 interaction (Isaacs & Trofimovich, 2012). Advanced listening skills play a key role in accessing aural L2 input available not only during actual L2 interaction but also via a variety of network media such as TV and Internet (Vandergrift, 2007), and extracting linguistic information from intake is hypothesized to be a first crucial step in SLA (Gass, 1997; VanPatten, 2004).

As Celce-Murcia et al. (2010) and Levis (2005) have pointed out, however, the teaching of pronunciation in L2 classrooms has been largely ignored over the past 20 years by the proponents of communicative language teaching. As a result, much pronunciation teaching is "still heavily influenced by commonsense intuitive notions" (Derwing & Munro, 2005, p. 380). In fact, pronunciation teaching has been notorious for its over-dependence on decontextualized practice, such as mechanical drills and repetition, reminiscent of the audio-lingual teaching methods of several decades ago (Trofimovich & Gatbonton, 2006). Previous intervention studies have indicated that the pedagogical potential of such focus-on-formS is limited to improvement only at a controlled level (see Saito, 2012 for research synthesis on 15 quasi-experimental studies published since 1990).

Furthermore, despite increasing research attention to the role of successful segmental perception in word recognition (Best & Tyler, 2007; Bundgaard-Nielsen et al., 2011a, 2011b; Cutler & Boersma, 2005) and its ultimate impact on holistic L2 comprehension (Stæhr, 2009), relevant findings have been restricted to the effects of laboratory training methods. During the training, L2 learners are intensively exposed to minimally-paired stimuli such as “rock”-“lock” produced by a number of native-speaking talkers (i.e., High Variability Phonetic Training) (for a review, see Thomson, 2012). The results of the previous studies have indicated that even adult L2 learners can alter their perception performance (e.g., Iverson, Hazan, & Bannister, 2005; Logan, Lively, & Pisoni, 1991) and transfer a gain to production levels (e.g., Bradlow et al., 1997). Notably, the pedagogical implications of these studies remain unclear, arguably because L2 learners listened to only a small number of phonetic targets and the length of instruction, in some cases, lasted for many hours (e.g., 15-22.5 hr in Bradlow et al., 1997). As a result, no single pronunciation or listening teaching study has ever been included in previous meta-analysis studies in the field of instructed SLA, due to the lack of sufficient discussion as to type of instruction (e.g., focus-on-form vs. focus-on-formS) and outcome measures (e.g., controlled vs. free constructed responses) (Norris & Ortega, 2000; Spada & Tomita, 2010).

FFI in SLA

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 grammar translation methods that introduce forms in a decontextualized manner, FFI is hypothesized to be most effective when implemented in content-based and communicative language classrooms in which conveying a meaningful message is a priority (Spada, 2011). Communicative focus on form in this way (a) helps L2 learners to develop their "form-meaning mappings" (Doughty, 2003; VanPatten, 2004) and (b) promotes a gradual transition from effortful to automatic use of rules (DeKeyser, 2007; Lyster, 2007). Specifically, Lyster and Ranta (Lyster, 2007; Ranta & Lyster, 2007) proposed a pedagogical sequence of FFI corresponding to differential levels of students' interlanguage development: Noticing → awareness → practice.

In FFI, teachers should first design a range of pre-planned tasks to promote learners' noticing of a target language feature in L2 input especially at the initial stage of interlanguage development (noticing phase) and then guide learners' analysis of the target feature with some degree of elaboration (awareness phase). Among effective FFI activities identified by researchers, those tested in this study include (a) structured input (i.e., learners are required to process linguistic form in input for meaning without being pressured to produce output: VanPatten, 2004), (b) typographically enhanced input (i.e., target structures are highlighted by means of emphatic stress or visual changes such as italics to induce learners to notice the forms in oral and written L2 input: Han, Park, & Combs, 2008), (c) focused tasks (i.e., learners are required to produce linguistically accurate output to successfully complete meaning-oriented tasks: Ellis, 2006). Finally, after learners successfully restructure and develop interlanguage representations, they are ready to be pushed to repetitively practice the target feature in production in communicatively authentic contexts (practice phase). At this stage, provision of outputprompting CF in a seemingly less planned fashion during online interaction could be a very effective technique.

With a general absence of research specifically investigating the pedagogical capabilities of FFI in L2 speech learning, Saito and Lyster (2012) took a first step towards testing the amenability of FFI techniques to teaching new sounds in conjunction with the *production* of /ɹ/ by Japanese learners. The results showed that providing CF (i.e., recasts) in response to learners' mispronunciation during FFI treatment played an important role in changing their L2 pronunciation performance. The relative importance of CF was attributed to its dual pedagogical function: Pronunciation-focused recasts provided students with pronunciation models while, at the same time, eliciting self-modified output.

Yet, that study had several methodological limitations. Furthermore, it resulted in more questions which future research needs to answer to obtain a better understanding of the relationship between FFI, CF, and L2 phonetic development. First, one could argue that such FFI effectiveness might be solely limited to production levels, arguably because adult L2 learners can sometimes carefully monitor their pronunciation forms without developing much perceptual awareness of the new sounds (Sheldon & Strange, 1982). Thus, the precursor study has yet to answer whether FFI can simultaneously impact both the *perception* and *production* domains (more similar to naturalistic L2 phonetic development: Flege, 2003) or if it simply reinforces learners' monitoring skills on their correct pronunciation forms of new sounds, regardless of the present state of their perceptual representations (Sheldon & Strange).

Second, the precursor study mainly measured the pedagogical potential of FFI only in the context of trained lexical items which appeared during instruction. It remains unclear to what extent learners can generalize their gain to novel lexical contexts beyond instructional materials (i.e., a transition from vocabulary to sound learning).

Finally, all participating students in the precursor study had been residing in an English speaking country (Canada) with much individual difference in their proficiency profiles—their length of residence (LOR) widely ranged from one month to 13 years (M = 15.5 months, SD = 31.8). Although the study found significant gains for the FFI and CF treatment, such results could have been confounded with the total amount and frequency of the participants' L2 use outside of the classroom in an English-as-a-Second-Language (ESL) setting. To isolate and re-examine the pure effects of FFI and CF on L2 speech learning, the study needs to be replicated, especially in conjunction with learners with homogeneous proficiency levels (e.g., LOR < 1 year) living in an English-as-a-Foreign-Language (EFL) environment, where they would be expected to have little exposure to spoken English outside of the classroom (for similar discussion on the differential effectiveness of phonetic training in ESL vs. EFL settings, see Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994).

Research Questions and Hypotheses

The current study was designed to examine which combination of FFI techniques (FFI with or without CF) impact the three different levels of L2 phonetic development of /ɹ/ by Japanese learners: (a) perception, (b) controlled production, (c) spontaneous production. It also aimed to investigate whether learners generalize the effects of FFI from trained to untrained lexical contexts, which might in turn provide evidence for their functional use of the lexical- and phonetic-level representational systems in a complementary fashion. The study's research questions and hypotheses are thus formulated as follows:

1. ***To what degree is FFI-only facilitative of L2 speech perception and production development of /ɹ/?***

Following Lyster and Ranta's FFI model, FFI-only in this stud is defined as providing a great deal of positive evidence of /ɹ/ through pre-planned noticing and awareness activities which included (a) structured input, (b) typographically enhanced input, and (c) focused tasks. It was predicted that such training without any further intervention on the part of the teacher, such as the provision of CF, would predispose the learners especially at the initial stage of interlanguage development to notice and practice /ɹ/ in a receptive manner. Because L2 pronunciation development requires not only learners' accurate perceptions of speech properties in L2 input (i.e., cognitive phase) and their actual usage of articulators to produce correct sounds (i.e., physical domain) (Flege, 2003), FFI-only might be insufficient to trigger learners' robust pronunciation improvement. It may be the case that some improvement would appear only at the perceptual level (i.e., perception precedes production, Flege, 1995, 2003) or, at best, may be transferred to the controlled-speech domain (Bradlow et al., 1997).

2. ***To what degree does adding CF to FFI increase the size of instructional gain?***

With respect to the efficacy of CF, previous descriptive research has generally shown that L2 learners tend to notice teachers' CF on their pronunciation errors and thus produce a great deal of modified output in various classroom settings (e.g., Sheen, 2006). Thus, the learners receiving CF during FFI are predicted to have more self-correction opportunities and occasion to proceduralize their productive use of the newly-acquired phonetic knowledge while comparing their nontarget pronunciation with the teacher's model pronunciation in contexts of interaction. Such communicative output practice is hypothesized to promote automatization of the L2 system

(Trofimovich & Gatbonton, 2006). Thus, learners will demonstrate gains not only at the perceptual level but also in both controlled and spontaneous production domains.

Method

Design

This study took place at a private language institute in Osaka, Japan, which granted access to its students, classrooms, and teachers for research purposes. After completing pre-tests, all students in the experimental groups first received four one-hour English Communication lessons over two weeks in which a range of FFI activities were embedded to encourage students to notice and practice the target pronunciation features of /ɹ/ (but without any explicit instruction) during meaningful discourse. For the FFI+CF group, instructors consistently provided CF on the mispronunciation of /ɹ/ while the FFI-only group did not receive any CF. Students in the control group received comparable communicative language instruction in terms of duration and content, but with no focus on phonetic form. The author observed the instruction from the back of the classroom to ensure consistency of treatments. All classes were video-recorded. Two weeks after the end of the lessons, all students completed post-tests and participated in a final interview.

Participants

The participants in this study consisted of (a) 49 Japanese learners of English; (b) two experienced native-speaking (NS) teachers; and (c) 10 NS baseline talkers.

Students. For recruitment purposes, the author advertised the English conversation lessons in the language institute and also in postsecondary educational institutions in the area of the research site, and on several social networking websites. Interested participants contacted the author and made appointments for their initial interviews and pre-tests. Of the 54 students who initially participated in this study, five failed to see the project through to completion for various personal reasons. The total number of students used in the final analyses was 49 (age: M = 30.5 years, SD = 10.1).

According to a preliminary language background questionnaire, although all learners had studied English formally for six years from Grades 7 to 12, typically through grammar translation methods, they reported that they had few opportunities to speak English, except for a few hours of conversation classes either at university or at language institutes. In addition, 33 of the 49 learners had never been abroad. Mean length of residence in an English-speaking environment was 7.9 months (SD = 21.4). Students were first randomly assigned to nine classes of six students each, and then each of the classes was assigned to one of three groups, with each group consisting of three classes. The three groups were (a) the FFI+CF group (three classes, n = 17), (b) the FFI-only group (three classes, n = 18), and (c) the control group (three classes, n = 14). The production (but not perception) data of the FFI+CF group and the control group were partially reported in another study (Saito, 2013a). In this study, however, the same production data was re-analyzed via new statistical and acoustic analyses in comparison with the FFI-only group (the main focus of this study is to examine the differential effectiveness of FFI with or without CF). Table 1 provides the details of the 49 learners according to three groups.

TABLE 1 AROUND HERE

Instructors. Two experienced NS instructors (one male from California, U.S., and one female from Ontario, Canada) who worked at the language institute participated in this study in the role of teacher. Whereas the first instructor taught one FFI+CF group and one FFI-only group, the second instructor taught two FFI+CF groups, two FFI-only groups, and three control groups.

NS baselines. Ten NS talkers (5 males, 5 females) participated by taking the same perception and production tests. They were undergraduate and graduate students at an English-speaking university in Montreal, Canada (age: $M = 21.5$, $SD = 1.6$). Their performance served as baseline data for comparison purposes.

FFI Treatment

Target words. In total, the target sound of /ɹ/ was embedded in 39 minimally paired words (including near minimally paired words; see Table 2) which frequently appeared in various FFI activities. All of these words were italicized and highlighted in red so that the learners might notice the target words in meaning-oriented lessons and become aware of the target feature of /ɹ/ at a lexical level (i.e., typographically enhanced input: Han et al., 2008).

TABLE 2 AROUND HERE

Instructional treatment. The main purpose of the meaning-oriented lessons was to teach English argumentative skills which entailed logical thinking, negotiation and debating, and public speaking abilities. To successfully complete each task, the learners were to accurately perceive and pronounce a few target words which were minimally paired with /l/. Among these activities were (a) argument critique, (b) English debating, (c) argument creation, and (d) public speaking.

1. *Argument Critique*: The students were given a set of logically problematic passages such as "***Ryan*** was able to drive his car well in the ***rain*** yesterday. So, he can drive his car in the snow without any problems," and taught how to find and critique these problems (i.e., analogy problem: "driving in the rain" ≠ "driving in the snow")

2. *Debating*: With critique techniques that they had learned, the students debated several topics such as "***Run***ning inside is better than ***run***ning outside" and "Is it good to have a ***rain***y day."
3. *Argument Creation and Public Speaking*: After learning how to use a logical sequence (i.e., Introduction → Evidence → Objection → Defense → Conclusion), the students were paired to develop their own arguments and make a public speech in front of the classes with good control of eye contact and a loud, clear voice. Topics used in this activity included "Is ***read***ing comic books good for children?" and "Is a sense of '***rat race***' among students harmful (e.g., tests, entrance examinations)?"

In addition to the main activities introduced above, three activities were used as "warm-up" games where students had to distinguish the target feature of English /ɹ/ from the Japanese tap /ɾ/ in perception and production to win the game. The instructors spent approximately 15 minutes on the games at the beginning of each lesson.

1. *English Karuta*: Thirty-six cards were placed on a table. Each card represented one lexical item, and had a relevant picture with the first letter of the word. When the teachers read a list of the words, the students tried to find and pick up the appropriate card as soon as possible. To get as many cards as possible, the students had to pay attention to the perceptual difference between /ɹ/ and /l/.
2. *English Card Game*: Each card had two identical sentences except one minimally paired word (e.g., "I found a beautiful ***leaf***" vs. "I found a beautiful ***reef***"), and the students in pairs took turns reading the sentences; the partner had to guess which sentence was read. To obtain several cards, the students had to differentiate their productions of the English /r-l/ contrast (there were 36 cards in total); the purpose of this activity was to promote

their awareness of their production of English /ɹ/-/l/ contrast but not their perceptual abilities (note that both students were Japanese learners of English).

3. *Guessing Game*: Each card had one vocabulary item written in standard orthography on the right-hand corner of the card. One learner would paraphrase the word, and the other would guess the vocabulary item his or her partner was trying to describe.

CF Procedure

For the FFI+CF group, the instructors consistently provided CF in the form of recasts in response to learners' mispronunciation or unclear pronunciation of /ɹ/ in the target words during the FFI treatment. Specifically, they were asked to recast only individual words that had been mispronounced, and to use a falling intonation when doing so (i.e., pronunciation-focused recasts). Such recast techniques were categorized as "partial recasts" and have been considered to be the most perceptually salient type of recast (Sheen, 2006).

To examine the relationship between recasts and self-modified output, the author carefully watched 12 hours of videotaped FFI+CF lessons (3 FFI+CF classes × 4 hr) and checked (a) the number of times the teachers provided recasts, and (b) to what degree they elicited students' repetition. Due to relative difficulty of accurately determining how successfully learners repaired their pronunciation errors during meaning-oriented classroom discourse (likely filled with unwanted noise), the author (a NS of Japanese) made a form of dichotomous coding on whether students made clear efforts to approximate English /ɹ/ (i.e., repair) or simply continued to substitute the Japanese tap /ɾ/ (i.e., needs-repair). For a similar judgement procedure on Japanese learners' /ɹ/ production in conversational speech and its rationale, see Riney, Ota, and Takata (2000).Three examples of the recast-repair sequence are as follows:

Example 1 for Repair (Argument Creation)

S: It might be better to let children read /ɾid/* comic books.

T: Read /ɹid/ ← RECAST

S: Read / ɹid/ comic books especially with adult guidance. ← REPAIR

Example 2 for Needs-Repair (English Debate)

S: We should avoid a wrong /ɾɔŋ/* sentence…

T: Wrong / /ɹɔŋ /. ← RECAST

S: Wrong /ɾɔŋ/ sentence. ← NEEDS-REPAIR

Example 3 for No Repair (English Debate)

S: If it rains /ɾeɪnz/*, we can…

T: Rains /ɹeɪnz/. ← RECAST

S: Yes. ← NO REPAIR

Control Group

The 14 learners in the control group received the same instruction on English argumentative skills as the other groups but without any FFI component on /ɹ/. All target words in the instructional materials were replaced with comparable words (“rain” → “typhoon,” “run” → “jog”). For the warm-up games, they did other communicative activities with no emphasis on pronunciation or listening practice.

Teacher Training

The researcher provided the instructors with four hours of training sessions over two days before the intervention commenced. First, they were given a package of instructional materials with a list of the target words. Next, they were told the content and purpose of each activity as well as the way they were to provide CF following students’ pronunciation errors.

Measures

The effects of FFI on L2 phonetic development were examined by means of a perception test, a controlled production test, and a spontaneous production test, including both trained and untrained lexical items. The testing sessions were individually conducted before and after instruction in a quiet room at the language institute. To avoid too much focus on form, especially in the spontaneous production task, the pre-post tests were conducted in the following order: the spontaneous production task → the controlled production task → the perception task.

Perception Tests

The two-alternative forced-choice identification task was used to measure the learners' perception performance of /ɹ/.

Materials. The test consisted of 70 test items containing 25 minimally-paired words beginning with word-initial /ɹ/ or /l/ ("rain" vs. "lane") together with 10 other minimally-paired words serving as distracters (e.g., "think" vs. "sink"). The minimally paired target words (*n* = 50) contrasted word-initial /ɹ/ and /l/ in three phonetic contexts: 20 consonant-vowel-consonant (CVC) singletons with front vowels, 10 CVC singletons with mid vowels, and 20 CVC singletons with back vowels (see Table 3).

TABLE 3 AROUND HERE

Three types of listening materials were prepared: (a) version A spoken by one male NS and one female NS of north-eastern American English; (b) version B spoken by the male NS instructor; and (c) version C by the female NS instructor. All stimuli were recorded by means of a Roland-05 Wave recorder in a quiet room at an English speaking university in Montreal (version A) or at the language institute (versions B and C). The tokens were digitized at a 44.1-

kHz sampling rate and normalized for peak intensity by means of speech analysis software, *Praat* (Boersma & Weenik, 2011).

Procedures. All test items were presented orthographically, and Japanese learners responded to all speech stimuli using an answer sheet. Several of the words were uncommon ones (e.g., lure, lope), and so the learners were told that all items were real minimally-paired words (and near minimally-paired ones) and were asked to focus on the contrasting sounds (e.g., /ɹ/-/l/, /θ/-/s/, /v/-/b/) rather than on lexical meaning. The assumption underlying the perception task was that, if the Japanese learners altered their L1-related perception patterns (i.e., sensitivity to F2 variation and transition duration) and started discriminating non-native phonetically-relevant acoustic parameters (i.e., sensitivity to F3 variation), they could successfully perceive /ɹ/ in the following manner: (a) identifying the sound with low F3 as /ɹ/ and (b) discriminating the sound with high F3 as /l/ (i.e., not /ɹ/).

Version A was used for all participants both at pre- and post-tests to measure any change in the perception level resulting from instruction. The participants also listened to Versions B or C at post-test sessions to see whether their instructional gain was restricted by training contexts (their teacher's voice).

Production Test

The two production tests (word reading and timed picture description) were designed to measure the learners' pronunciation performance of /ɹ/ at two different levels of processing (controlled- vs. spontaneous-speech levels). Their speech tokens were recorded by means of Roland-05 Wave recorder, at 44.1 kHz sampling rate and a 16-bit resolution and a unidirectional microphone (DM-20SL).

Controlled production. To elicit controlled production of /ɹ/, the learners read a list of 40 words of which 15 were target words. The test tokens were CVC singletons except "Ryan" (CVVC), including six tokens with front vowels, three tokens with central vowels, and six tokens with back vowels. According to Vocabulary Profiler (Cobb, 2011), these tokens fell within the first 2000 word families except "Ryan" and "rink" (the test items are relatively high frequency words). These 15 words included (a) 10 trained lexical items and (b) five untrained lexical items. The test tokens are identified in Table 4.

Spontaneous production. As a reliable way to measure spontaneous production abilities of certain linguistic features, SLA researchers have emphasized the importance of eliciting learners' use of language for some communicative purpose (i.e., they are required to pay equal attention to grammatical, phonetic, lexical, and pragmatic aspects of language) (Spada & Tomita, 2010) and under time pressure (i.e., they are not given much planning time to access their explicit knowledge) (Ellis, 2005). Among such cognitively-demanding tasks is a picture description task wherein learners are guided to use certain linguistic structures with written or oral prompts while describing pictures; this task has been widely used in previous instructed SLA research (for a summary, see Ellis, 2002). In L2 phonology, it has been found that the picture description task induces more segmental errors as well as fluency problems than controlled production tests such as word and sentence reading tasks (Derwing et al., 2004; Lin, 2003; Rau et al., 2009). To elicit spontaneous production of /ɹ/ by Japanese learners, a timed picture description task was adapted in this study as follows.

First, 16 pictures including six distracters were prepared, 10 of which were designed to lead the learners to pronounce one target word including word-initial /ɹ/. These tokens were CVC singletons, including four tokens with front vowels, two tokens with central vowels, and four tokens with back vowels (5 trained items + 5 untrained items). According to Vocabulary Profiler (Cobb, 2011), the target lexical items fell within the first 2000 word families except “route” (which was included in the academic word list). These words are listed in Table 4. Subsequently, the learners were asked to describe pictures one after another under communicative pressure:

1. Learners were given 10 seconds to memorize a list of four key words.
2. Right after the list was taken away, the learners were asked to describe two pictures in a row without any planning time by using two key words for each picture; one of them was a target word including /ɹ/ at a word-initial position.
3. After describing the pictures, they moved on to the next four key words for another set of two pictures.

TABLE 4 AROUND HERE

Acoustic Analysis

In accordance with the acoustic analysis performed for natural speech tokens of word-initial /ɹ/ established by Flege (Flege, Takagi, & Mann, 1995b), F3, F2, F1, and the transition duration of all speech tokens were measured through linear predictive coding spectra by means of *Praat* (Boersma & Weenik, 2011). The beginning of word-initial English /ɹ/ was first identified via both the spectrographic representations and wave forms of the speech tokens in conjunction with the onset of the energy for all three formants. The cursor was put at this point to measure F3, F2 and F1 in Hz. With respect to /ɹ/ embedded in continuous speech (i.e., the spontaneous production tokens), the severe dip in F3 (local peak) was carefully identified, given that English /ɹ/ exhibits relatively low F3 values compared to other vowel and consonant sounds in the English phonetic system. Finally, transition duration was measured by dragging a cursor from the beginning point of F1 transition to the endpoint of the F1 or F3 transition (Hattori & Iverson, 2009).

Normalization

Given that spectral information (i.e., F3, F2, F1 values) significantly varies due to anatomical differences in individual vocal tract length, raw acoustic values were adjusted in the following normalization procedure (for a more detailed account of this procedure, see Lee, Guion, & Harada, 2006; Yang, 1996). First, a mean F3 value of /æ/ elicited from 10 monosyllabic words in the controlled production test (i.e., *fan, tap, map, bat, mad, cap, bag, lab, cat, dad*) was calculated for each participant (N = 49 Japanese learners + 10 NS baseline speakers). One female English speaker was randomly selected as a reference, and her mean F3 value (i.e., 2897 Hz) was divided by those of the other participants to provide their own k factors. Then, all formant values (F3, F2, F1) of /ɹ/ for each participant were multiplied by the individual respective k factor. Finally, all raw values in Hz were converted to Bark to reduce the nonlinear relationship between the formant frequencies and the corresponding perceived semivowel quality (more similar to human perception of /ɹ/).

Teacher Evaluation

To examine the extent to which the students' improvement resulting from instruction could be perceptible, the two instructors rated the quality of each students' /ɹ/ on the first and last day of lessons (the male NS teacher A for one FFI-only and one FFI+CF class; the female NS teacher for two FFI-only classes, two FFI+CF classes, three Control classes). This decision reflects L2 classroom practice: Teachers not only deliver instruction but also provide adequate assessments on their students' performance at regular intervals to guide and motivate them towards continued L2 development inside and outside the classroom.

First, a descriptor of a 9-point scale was adapted and modified from Flege et al.'s (1995) 6-point scale, and the rating criteria were explained as follows: 1 (*very good /ɹ/*) → 2 (*good /ɹ/*) → 3 (*probably /ɹ/*) → 4 (*possibly /ɹ/*) → 5 (*neutral exemplars, neither /ɹ/ nor /l/*) → 6 (*possibly /l/*) → 7 (*probably /l/*) → 8 (*good /l/*) → 9 (*very good /l/*). To enhance the validity of teacher evaluation, we asked the instructors to not only (a) assess each of their students as objectively as possible, but also (b) practice the current listening procedure prior to their actual teaching assignment, by listening to and evaluating the total number of 50 /ɹ/ tokens randomly selected from the data pool of the precursor study (Saito & Lyster, 2012) whereby 65 Japanese learners read singletons including word-initial /ɹ/ via word and sentence reading and picture description tasks.

Results of Perception Tests

To avoid ceiling effects, the four learners (n = 1 from FFI-only, n = 1 from FFI+CF, n = 2 from Control) whose listening scores were more than 90% at pre-tests ($x \geq 45$ out of 50 points) were eliminated from the final analyses. For all following inferential statistical analyses, an alpha level was set at a $p < .05$ level. Cohen's d was also calculated to measure the magnitude of instructional effectiveness between two contrast groups of means.[1]

Initial Performance

The 45 learners in the experimental and control groups took version A (spoken by untrained male and female talkers) at pre-tests, and scored 63.4% (31.70 points out of 50 items) on average. A one-way ANOVA did not yield any significant group difference among the three

[1] According to Oswald and Plonsky (2010), effect sizes in the field of instructed SLA are roughly classified as small ($d < 0.70$), medium ($0.70 \leq d < 1.00$), or large ($1.00 \leq d$).

groups at pre-tests, F (2, 42) = 2.493, p = .095. The 10 NSs completed test versions A, B, and C to ascertain the construct validity of the listening stimuli. Not surprisingly, they demonstrated perfect scores for all test materials (M = 50 points).

Pre-Post Test Performance

The learners took not only version A but also version B or C (spoken by their instructors) at post-tests. Their raw listening scores are plotted in Figure 1. According to the inferential statistics described below, FFI-only group significantly outperformed Control, while FFI+CF did not.

ANCOVAs were conducted on their post-test listening scores yielded by both version A and versions B or C to investigate any group difference resulting from instruction at the time of post-tests, with their initial perception abilities of /ɹ/ (measured via version A at pre-tests) as a covariate. Whereas the ANCOVA did not yield a significant group difference in the case of a trained talker (version B or C), F (2, 41) = 1.764, p = .184, it revealed a significant effect for Group in the case of an untrained talker voice (version A), F (2, 41) = 3.801, p = .031. Bonferroni post-hoc comparisons, performed on adjusted means, found a significant difference between the FFI-only group (M = 35.59 points) and the control group (M = 31.25 points). The effect size analysis revealed a medium effect size for FFI-only (d = 0.87).

FIGURE 1 AROUND HERE

To further examine the influence of lexical contexts on these gains, ANCOVAs were conducted separately on learners' post-test performance on version A according to trained lexical items and untrained lexical items. In sum, the FFI-only group significantly improved their perceptual abilities of /ɹ/ especially in untrained lexical items produced by untrained talkers.

For trained items, the ANCOVAs yielded marginally significant effects for Group, F (2, 41) = 3.227, p = .050. Bonferroni post-hoc comparisons found a significant difference between the FFI-only group (M = 21.42 out of 30 points) and the control group (M = 19.07 points) with medium effects (d = 0.71). For untrained items, the ANCOVAs revealed a significant Group effect, F (2, 41) = 3.370, p = .004. The post-hoc comparisons showed that the FFI-only group (M = 14.49 out of 20 points) outperformed the control group (M = 11.80 points) with large effects (d = 1.08).

Results of Production Tests

Given the pattern of distinct formant distributions between /ɹ/ and Japanese tap /ɾ/ discussed earlier (Hattori & Iverson, 2009; Saito & Brajot, 2013), the following benchmark (Japanese tap /ɾ/ → English /ɹ/) was used to interpret the acoustic results:

- F3: 14.10-15.70 Bark (2400-3000 Hz) → 11.40-12.60 Bark (1600-1900 Hz)
- F2: 11.80-13.20 Bark (1700-2100 Hz) → 7.90-11.00 Bark (900-1500 Hz)
- Transition duration: 5-20ms →50-100ms

In the following analyses, compared to the Japanese tap, the degree of reduction in F3 (the *primary* cue) is considered a valid index of "extent of acquisition," and the degree of change in F1, F2 and transition duration (the *minor* cues) an index of "use of the interlanguage strategy."

Initial Performance

The descriptive results of the acoustic values are summarized in Table 5. According to the inferential statistics described below, the Japanese learners in this study tended to substitute the Japanese tap for English /ɹ/ at the beginning of the project.

First, we examine how the three Japanese groups (Control, FFI-only, FFI+CF) produced /ɹ/ under various levels of processing (controlled vs. spontaneous) and lexical contexts (trained vs. untrained items), compared to a native English-speaking group (NS baseline). Their four acoustic values of /ɹ/ (F3, F2, F1, transition duration) were separately submitted to three-way ANOVAs with one between-group factor (4 groups) and two repeated factors (2 task conditions [controlled, spontaneous] × 2 lexical conditions [trained, untrained lexes]). Main effects of Group were found for (a) F3, F (3, 55) = 49.107, $p < .001$; (b) F2, F (3, 55) = 7.88, $p < .001$; and (c) transition duration, F (3, 55) = 30.98, $p < .001$. The results of Bonferroni post-hoc comparisons showed that the Japanese learners' /ɹ/ production at post-tests was significantly different from the NS baseline, across various lexical and task conditions, with respect to higher F3 (14.40-14.70 Bark), higher F2 (11.10-11.90 Bark) and a shorter transition duration (25-35ms) ($p < .001$).

TABLE 5 AROUND HERE

Pre-post Test Performance

Primary Cue (F3). The descriptive results of F3 values are plotted in Figure 2. A visual inspection of the figure indicates that both FFI-only and FFI+CF groups generally reduced F3 values from the range of the Japanese tap (F3 > 14.10 Bark) to English /ɹ/ (F3 < 14.10 Bark) in all lexical and task conditions. Yet, their post-test performance (F3 = 13.50-14.50 Bark) was still substantially different from the range of the NS baseline (F3 = 12.00-12.30 Bark).

FIGURE 2 AROUND HERE

To measure the impact of instruction on the acquisition of F3, the F3 values were submitted to a four-way ANOVA with one between-group factor (3 groups [FFI+CF, FFI-only, Control]) and three repeated factors (2 task conditions [controlled, spontaneous] × 2 lexical conditions [trained, untrained lexes] × 2 time conditions [pre-, post-tests]). The results showed a significant Group × Time interaction effect, F (2, 46) = 10.489, $p < .001$. According to Bonferroni pairwise comparisons, both the FFI+CF and FFI-only groups produced significantly lower F3 values (more targetlike exemplars of /ɹ/) than the control group at post-tests ($p = .042$, $d = 0.88$ for FFI+CF, and $p = .006$, $d = 1.08$ for FFI-only).

Secondary cues (F2, F1, transition duration). To investigate how FFI influenced the use of the interlanguage strategy (change in F2, F1, transition duration), the following subsections present the results of descriptive and inferential statics for each acoustic dimension.

- ***F2***: The descriptive results of F2 values are plotted in Figure 3. Inspection of the figure reveals that the experimental groups equally reduced F2 values (i.e., producing /ɹ/ with more tongue retraction). Given the range of NS baseline in Table 5 (F2 = 10.00-10.40 Bark), the figure suggests that FFI led the learners to reach the range of NS baseline across various lexical and task contexts.

FIGURE 3 AROUND HERE

A four-way ANOVA (Group × Task × Lexis × Time) yielded a significant interaction effect of Group and Time, F (2, 46) = 5.428, $p = .008$. Bonferroni pairwise comparisons showed that the FFI+CF and FFI-only groups outperformed the control group at post-tests with significantly lower F2 (p = .006, d = 0.46 for FFI+CF, and p = .030, d = 0.83 for FFI-only).

- ***F1***: The descriptive results of F3 values are visually plotted in Figure 4. Although the F1 range does not differ between the Japanese tap and English /ɹ/ (F1 = 600-900 Hz) (Hattori & Iverson, 2009) and thus the Japanese learners do not need to modify F1 to produce /ɹ/, a visual inspection of the figure shows that all Japanese learners generally reduced F1 (i.e., producing /ɹ/ with higher tongue positions).

FIGURE 4 AROUND HERE

A four-way ANOVA (Group × Task × Lexis × Time) indeed yielded only a significant main effect of Time, F (1, 46) = 26.123, p < .001. That is, all Japanese learners produced /ɹ/ with lower F1 values at post-tests than pre-tests (M = 4.37 Bark → 4.09 Bark) with small effects (d = 0.46).

- ***Transition duration***: The descriptive results of transition duration are plotted in Figure 5. Inspection of the figure indicates that the Japanese learners in the experimental groups tended to lengthen the duration of /ɹ/ from 30 ms (close to the Japanese tap: 5-20 ms) to 50 ms (close to English /ɹ/: 50-100ms). Given the range of the NS baseline in Table 6 (60-90ms), the figure suggests that FFI led the learners to reach nativelike temporal aspects of English /ɹ/.

FIGURE 5 AROUND HERE

A four-way ANOVA (Group × Task × Lexis × Time) found a significant interaction effect of Group and Time, F (2, 46) = 8.675, p = .001. According to Bonferroni post-hoc comparisons, both of the experimental groups (FFI+CF, FFI-only) produced /ɹ/ with significantly longer transition durations than the control group at post-tests (p = .013, d = 0.96 for FFI+CF, and p = .008, d = 1.12 for FFI-only).

Teacher Evaluation

Validity. The rating of the subset of /ɹ/ tokens (n = 50) demonstrated a high correlation between the two instructors, r (49) = .855, p < .001 without significant difference between their actual rating scores, F (1, 98) = .516, p = .474. Since their judgement patterns were highly comparable, their scores were combined.

Pre-post. The instructors' 9-point rating scores (1. *very good /ɹ/* - 9. *very good /l/*) were plotted in Figure 6. According to two-way ANOVA (Group × Time), the results yielded a significant interaction effect, F (2, 46) = 14.206, p < .001. Bonferroni pos-hoc comparisons showed that, although both teachers generally judged their students' /ɹ/ production on Day 1 as "neutral exemplars" (M = 5.10), a significant improvement over Time was found for the FFI+CF group (M = 5.89 → 3.94, p < .001, d = 1.00) the FFI-only group (M = 5.12 → 4.41, p = .005, d = 0.39).

FIGURE 6 AROUND HERE

Recast-repair Analyses

Eighteen learners in the FFI+CF group received 579 recasts and repeated 474 of them, yielding a relatively high repetition rate (81.2%). On average, each learner in the FFI+CF group received 32.17 recasts (*M* = 8.0 per lesson). In line with the previous descriptive studies (e.g., Sheen, 2006), the results showed that the learners likely noticed the instructors' corrective intention of pronunciation-focused recasts, and they were thus pushed to process a great number of enhanced output opportunities during FFI activities (i.e., repetition of recasts). With respect to the quality of repetition, however, the students successfully repaired only 59 out of 579 recasts, with a relatively low repair rate (10.2%). Although many of the participating students repeated the teachers' recasts, they actually did so simply by substituting the Japanese tap /ɾ/, instead of making a clear effort to approximate English /ɹ/.

Discussion

The main goal of the current study was to investigate how two types of FFI—FFI with and without CF—could promote the acquisition of /ɹ/ by Japanese learners in EFL settings. FFI effectiveness was assessed via three outcome measures (perception, controlled production, spontaneous production) and also according to two lexical contexts (trained, untrained items). According to the pre-post test scores, the learners in the FFI+CF and FFI-only groups revealed slightly different patterns in instructional gains. These results of this study will be further discussed below in conjunction with previous findings in the field of instructed SLA and L2 phonology in an interdisciplinary manner.

Effects of FFI

Unlike decontextualized pronunciation drills and intensive laboratory perception training methods, this study aimed to test a combination of communicative input and output activities (FFI) in L2 segmental learning. With respect to the perception measures, the results showed that the FFI-only group outperformed the control group at the time of post-tests, with instructional gains particularly apparent in the context of the novel lexical items spoken by the untrained talkers. The amount of improvement resulting from four hours of FFI (5.9% gain) could be considered as medium-to-large, and comparable to other intensive lab training studies (e.g., Logan et al., 1991 for 8% gain after 10 hr of training). With respect to the production measures, the results of the acoustic analyses showed that the FFI-only group not only enhanced their use of an interlanguage strategy (F2 reduction for tongue retraction and increase in transition duration for lengthening the phoneme) but also developed new articulatory configurations for English /ɹ/ (i.e., F3 reduction for labial, palatal and pharyngeal constrictions) to approximate a more nativelike production of English /ɹ/. Importantly, the learners could also generalize the instructional gain from trained to untrained lexical contexts as well as from controlled to spontaneous speech contexts. This suggests that these learners developed a robust phonetic category of /ɹ/ with various levels of processing abilities. Similar to the acoustic data, the two NS instructors perceived their students' general improvement from neutral to intelligible exemplars of /ɹ/ between the first and last day of the lessons. As expected earlier, the efficacy of FFI on the development of L2 speech perception and production could be due to several factors.

First, FFI could successfully draw L2 learners' selective attention towards phonetic form, while maintaining their primary focus on meaning. Different from focus-on-formS interventions, repetitive practice of language under communicatively authentic contexts is believed to encourage students to appropriately transfer what they have learned in classrooms to the world outside of the classrooms (Lyster, 2007; Trofimovich & Gatbonton, 2006). Second, teacher talk in such meaning-oriented classrooms could provide optimal linguistic input to facilitate L2 phonetic development, such as higher pitch, simplified prosody, and vowel and consonant hyperarticulation. Many speech researchers have claimed that the acoustic enhancements could be the best way for adult L2 learners to circumvent L1 interference effects and increase their awareness of new sounds (see Kuhl, 2000, p. 11855). In fact, Iverson et al. (2005) found that exposing Japanese learners to synthetically exaggerated speech stimuli, with F3 and F2 variation maximized and duration lengthened, helped them improve their perception abilities of the non-native /ɹ/-/l/ contrast (see also McCandliss et al., 2002).

Finally, the importance of social interaction may play a crucial role in successful L2 speech learning processes. In early bilingual literature, it has been well documented that speech learning tends to occur preferentially for signals from a live person rather than those from televised or audio-only presentations. For instance, Kuhl, Tsao, and Liu (2003) found that English infants who interacted with four native Mandarin caretakers for approximately six hours demonstrated remarkable L2 perception learning for Mandarin phonetic contracts, although the infants who listened to the same speech stimuli but via televised or audio-only presentation failed to do so. As Kuhl (2007) argued, "social interaction might be effective *because* it involves other humans, or because features inherent in social settings, such as interactivity and contingency, are critical for learning" (p. 115, *emphasis original*). In the adult bilingual literature, this study investigated in depth the effectiveness of social interaction on speech learning processes by adopting a FFI research framework into pronunciation and listening teaching contexts. Our results provided empirical evidence that processing new sounds through online interaction with teachers could be facilitative of various aspects of L2 phonetic development even within a limited amount of instructional treatment (i.e., 4 hr).

Adding CF to FFI

As we confirmed the effectiveness of FFI itself on the development of speech perception and production of /ɹ/, we now turn our discussion to the issue of whether and to what degree adding CF to FFI could further increase (or decrease) the size of the gain. In our research design, CF was operationalized as pronunciation-focused recasts (teachers recasted students' mispronunciation of /ɹ/ during FFI activities), and it served as production enhancements in light of the relatively high rate of repetition (students perceived almost all recasts as correction and they repeated the teacher's model pronunciation). Similar to the precursor study (Saito & Lyster, 2012), the FFI+CF group showed significant improvement in their production abilities of /ɹ/. Notably, however, they did not show any significant gain in the perception measures. That is, the findings in this study suggest that (a) FFI itself can sufficiently promote the acquisition of /ɹ/, but (b) the acquisitional value of CF in L2 speech learning remains unclear. These unexpected findings (i.e., FFI-only = FFI+CF) need to be carefully discussed along with the various possible interpretations.

First of all, the findings in the current study shed light on the role of learners' interlanguage proficiency levels in recast effectiveness. In L2 morphosyntax development, previous research has consistently shown that recasts facilitate acquisition when learners have a sufficient amount of knowledge for target forms but not when learners are considered at lower stages of development. Based on the well-established developmental sequence of English question formation, Mackey and Philp (1998) found that recasts positively influenced learners who were developmentally ready to acquire the target feature (Stage 3), but not those who lacked developmental readiness (Stage 2) (see also Ammar & Spada, 2006 for the acquisition of possessive determiners). In their review of L2 recast research, Nicholas, Lightbown and Spada (2001) pointed out, "recasts can be effective if the learner has already begun to use a particular linguistic feature and is in a position to choose between linguistic alternatives" (p. 752). These findings in turn indicate that recasts help advanced learners consolidate their already-acquired knowledge rather than helping beginner learners acquire completely new knowledge.

Importantly, the analyses of the pre-test performance showed that most of the learners in this study substituted the Japanese tap /ɾ/ for English /ɹ/ both in perception and production. Furthermore, the post-hoc analysis demonstrated that despite their high noticing and awareness of teachers' recasts during FFI lessons (i.e., 81.2% of repetition rate), many learners in the FFI+CF group continued to use the Japanese tap /ɾ/ to repeat received recasts (i.e., 10.2% of repair rate). This suggests that (a) the participating students had yet to develop or establish any robust mental representations of /ɹ/ at the onset of the project; and thus (b) such beginner learners were unlikely to benefit from CF treatment, arguably because they lacked the phonetic knowledge (i.e., how to repair their mispronunciation of /ɹ/) necessary to process pronunciation-focused recasts.

Another crucial index of interphonology proficiency is the amount of L2 experience (usually measured as LOR in a target language country). Previous studies have shown that the first few months of intensive exposure to the L2 tend to make a great impact on the developing L2 phonetic system (for a review, see Piske, MacKay, & Flege, 2001). Although the learners in this study did report some prior experience in L2 (LOR: *M* = 3.4 months for FFI+CF, *M* = 7.6 months for FFI-only)[2], most of them (33 out of 49 learners: see Table 1) had never been abroad and, importantly, they were in EFL settings with only minimal use of English in classroom settings at the time of the project (a few hours per week). Recent psycholinguistic SLA studies have shown that adult SLA in foreign classroom settings without many opportunities to use language for meaning does not necessarily promote the integration of explicit knowledge into their mental representational system (e.g., Jian, 2007).

[2] LOR for FFI-only was almost double that for FFI+CF, mainly because one learner in the former group noted outstandingly long LOR (eight years) in Canada.

Taken together, it is possible that these learners, with little or no integrated phonetic knowledge in their L2 developing system, could be developmentally “unready” to benefit from recasts and concomitant opportunities for output practice. As a sequence for L2 phonetic development, many researchers have claimed that phonetic categories are based in perception, which in turn indicates that, once L2 learners hear perceptual aspects of new segmental sounds and dissimilate them from L1 counterparts, they start internalizing their newly-acquired phonetic knowledge as mental representations (Flege, 1995, 2003; Kuhl, 2000). In this respect, the beginner should be encouraged to notice and practice the target feature of /ɹ/ only through FFI in a receptive mode without much pressure for modified output (recast treatment).

Similarly, the importance of the input-driven view of SLA (receptive learning without much pressure for production) especially for the early stage of interlanguage development has also been extensively discussed in the domain of L2 morphosyntax studies (for review, Gass, 1997). For example, VanPatten (2004) maintains that L2 learners first need to extract morphosyntactic information from L2 lexical input in a receptive mode. This is believed to be a first step towards enhancing the quality of the L2 developing system, which is responsible for their performance in output at a later stage of SLA processes (i.e., a comprehension-first view). Following his theoretical model of SLA, VanPatten and his colleagues conducted a series of empirical studies, providing some evidence that beginning L2 learners who received only input-based practice in the form of processing instruction without any pressure to actually produce L2 output improved not only their comprehension abilities but also production abilities (e.g., VanPatten, 2004; VanPatten & Cardierno, 1993).

The question is, then, when should pronunciation-focused recasts as production enhancements be introduced to facilitate L2 speech learning? Given that many SLA researchers have argued that the primary pedagogical function of CF (production enhancement) is to automatize the retrieval of existing knowledge (Lyster, 2007; Nicholas et al., 2001), we suggest ensuring that L2 learners have already established their increasing awareness towards a new sound to some degree in order to make the best use of subsequent recast treatment during FFI activities.

As a pedagogic means to increase less-advanced learners' attention to phonetic form, it would be intriguing to provide explicit phonetic information at the beginning of FFI lessons. In Saito's (2013a, 2013b) experiment, for example, teachers exaggerated their pronunciation of /ɹ/ and provided explanation on the relevant articulatory configurations (i.e., labial, palatal, and pharyngeal constrictions). In this way, the teachers were able to help their students decode and create new perceptual categories for L2 segmentals from ambient language input, as well as activate the correct use of articulators to pronounce these sounds under little communicative pressure before they moved onto practicing FFI activities. As a result, not only were those receiving explicit phonetic information able to make a substantially large change on the primary acoustic property of their English /ɹ/ production (i.e., F3) (Saito, 2013a), but their improvement was also perceptible to human listeners as well as apparent in the perception domain (Saito, 2013b).

Another remedial strategy could be the adoption of different types of CF, such as metalinguistic correction, i.e. teachers providing explicit information while at the same time reformulating students' linguistic errors during FFI lessons. In L2 morphosyntactic development, for example, Sheen (2007) found that metalinguistic correction significantly facilitated the acquisition of English articles by adult ESL learners, although recasts alone did not. In this vein, future studies are needed to examine how we can further enhance the magnitude of recast effectiveness by including other mediating FFI techniques (e.g., explicit phonetic information, metalinguistic correction).

Limitations

Given that the current study took an exploratory approach to examining the effects of FFI on L2 speech learning, certain shortcomings must be acknowledged with an eye toward future replication. First, although the two-alternative forced-choice identification task was the only perception measure in this study, the instructional benefits on L2 perception need to be scrutinized via a range of perception tests in experimental phonetics, such as identification and discrimination tasks with the synthetic F2 × F3 continua (Iverson et al., 2003), in noisy conditions (Munro, 1998), and with reaction time instruments (Lively et al., 1994). Second, given that previous studies showed that L2 segmental perception is subject to frequency and familiarity of test words (Flege, Takagi, & Mann, 1996), future studies should carefully control this variable to investigate in depth the FFI effectiveness on L2 perception development. Last, L2 production development was measured based on acoustic analyses and teacher evaluation. Whereas the teachers' judgement could be considered as reliable in conjunction with their accumulated teaching experience in EFL classrooms, their evaluation could also easily become lenient or biased due to their actual interaction with the students. Although the experimental groups reached the nativelike range for minor cues (F2, F1, transition duration) after receiving FFI, we need to acknowledge that the impact of FFI on developing the new phonetic representation (F3) was still not substantial (their post-test performance of F3 [13.50-14.50 Bark] was still considerably different from the nativelike range of English /ɹ/ [12.00-12.30 Bark]). So, future research could implement human rating sessions to see how such changes in F3, F2, F1 and transition duration can ultimately impact the perception of *naïve* listeners who have little familiarity with Japanese-accented English (cf. Saito, 2013b).

Conclusion

Building on the instructed FFI research framework, this study examined in depth how two types of FFI (FFI with or without CF) can facilitate the acquisition of /ɹ/ by adult Japanese learners, and help them improve three different levels of L2 performance (perception → controlled production → spontaneous production) across two lexical contexts (trained, untrained items). The results showed communicative focus on form via FFI can be facilitative of the development of L2 speech perception and production of /ɹ/. The FFI effectiveness could be tied to the fact that L2 learners are to repetitively practice phonetic forms through social interaction with their teachers under communicatively authentic conditions. Although the learners in this study showed high levels of noticing, the provision of CF as a form of recasts did not appear to be effective for acquisition, arguably because recasts help L2 learners increase control over their existing knowledge rather than acquire completely new knowledge. In fact, the EFL Japanese learners in this study had minimal opportunities to use English on a daily basis and had yet to have much phonetic knowledge of the target sound /ɹ/ at the time of the project. Thus, the results in turn suggest that these beginners need to focus more on noticing the perceptual aspects of new sounds in a receptive mode in order to start establishing relevant phonetic representations of the sounds before processing production enhancements, such as pronunciation-focused recasts. To further examine the role of learners' levels of representation and processing abilities in CF effectiveness, we therefore make a strong call for more instructed L2 speech research of this kind, specifically research which recruits a range of L2 learners with different proficiency profiles. Such future studies will as a result shed some light on when and how recasts can help, in particular with respect to how advanced and intermediate learners (or those who receive explicit information before or during FFI lessons) can enhance their perceptive and productive performance of new L2 sounds.

References

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition, 28*, 543-574.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception. In O. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Boersma, P., & Weenik, D. (2011). *Praat: Doing phonetics by computer*. Retrieved from http://www.praat.org (Last viewed September, 2011).

Bradlow, A. R. (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. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors. *Journal of the Acoustical Society of America, 106*, 2074-2085.

Bradlow, A., Pisoni, D., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /ɹ/ and /l/. *Journal of the Acoustical Society of America, 101*, 2299–2310.

Bundgaard-Nielsen, R., Best, C., & Tyler, M. (2011a). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition, 33*, 433-461.

Bundgaard-Nielsen, R., Best, C., & Tyler, M. (2011b). Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. *Applied Psycholinguistics, 32,* 51-67.

Celce-Murcia, M., Brinton, D., Goodwin, J., & Griner, B. (2010). *Teaching pronunciation*. Cambridge: Cambridge University Press.

Cobb, T. (2011). *The compleat lexical tutor*. Retrieved from http://www.lextutor.ca/vp/

Cutler, A., & Broersma, M. (2005). Phonetic precision in listening. In W. J. Hardcastle, & J. Mackenzie Beck (Eds.), A figure of speech: A Festschrift for John Laver (pp. 63–91). London: Lawrence Earlbaum.

DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge, UK: Cambridge University Press.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research–based approach. *TESOL Quarterly*, *39*, 379–397.

Derwing, T.M., Rossiter, M.J., Munro, M.J. & Thomson, R.I. (2004). L2 fluency: Judgments on different tasks. *Language Learning, 54,* 655-679.

Ellis, R. (2001). Investigating form-focused instruction. *Language Learning, 51*, 1–46.

Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? *Studies in Second Language Acquisition, 24*, 223–236.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a Second Language: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. *AILA Review, 19*, 18-41.

Espy-Wilson, C., Boyce, S., Jackson, M., Narayanan, S., & Alwan, A., (2000). Acoustic modeling of American English /ɹ/. *Journal of the Acoustical Society of America*, *108*, 343-356.

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

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J., Frieda, E., Walley, A., & Randazza, L. (1998). Lexical factors and segmental accuracy in second-language speech production. *Studies in Second Language Acquisition, 20*, 155-188.

Flege, J., Munro, M, & MacKay, I. R. A. (1995a). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Flege, J., Takagi, N., & Mann, V. (1995b). Japanese adults learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38*, 25-55.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of Acoustical Society of America, 99,* 1161-1173.

Gass, S. (1997). *Input, interaction, and the second language learner.* Mahwah, NJ: Erlbaum.

Guion, S., Flege, J., Ahahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*, 2711-2725.

Han, Z., Park, E., & Combs, C. (2008). Textual enhancement of input: Issues and possibility. *Applied Linguistics, 29*, 597-618.

Hattori, K., & Iverson, P. (2009). English /ɹ/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America, 125*, 469-479.

Imai, S., Walley, A., & Flege, J. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish–accented words by native English and Spanish listeners. *Acoustical Society of America*, *117*, 896–907.

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, *34(4)*.

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

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, 847-857.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language acquisition. *Language Learning, 57*, 1-33.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science, 97,* 11850-11857.

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science, 10*, 110-120.

Kuhl, P. K., Tsao. F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*, 9096-9101.

Lee, B., Guion, S.G., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition 28*, 487-513.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*, 367–377.

Lin, Y. (2003). Interphonology variability: sociolinguistic factors affecting L2 simplification strategies. *Applied Linguistics, 24*, 439–464.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /ɹ/and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, 96, 2076-2087.

Logan, J., Lively, S., & Pisoni, D. (1992). Training Japanese listeners to identify English /ɹ/ and /l/: A first report. *Journal of the Acoustical Society of America, 89*, 874-886.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam/Philadelphia: John Benjamins.

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.

McCandliss, B., Fiez, J., Protopapas, A., Conway, M., & McClelland, J. (2002). Teaching the [r]-[l] discrimination to Japanese adults: Behavioral and neural aspects. *Physiology & Behavior*, *77*, 657–662.

Munro, M. J. (1998).The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, 20, 139-154.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, *51*, 719-758.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50,* 417-528.

Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second-language research: Choices and challenges. *Annual Review of Applied Linguistics, 30,* 85-110.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accents in an L2: a review. *Journal of Phonetics*, *29*, 191–215.

Ranta, L., & Lyster, R. (2007). a cognitive approach to improving immersion students' oral language abilities: The Awareness-Practice-Feedback sequence. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 141-160). Cambridge: Cambridge University Press.

Rau, D., Chang, A., & Tarone, E. (2009). Think or sink: Chinese learners' acquisition of the voiceless interdental fricative. *Language Learning*, *59*, 581-621.

Riney, T., Takada, M., & Ota, M. (2000). Segmentals and global foreign accent: The Japanese flap in EFL. *TESOL Quarterly, 34*, 711–737.

Rose, K., & Kasper, G. (Eds). (2002). *Pragmatics in language teaching*. Cambridge: CUP.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly,* 842-854.

Saito, K. (in press, 2013a). Re-examining effects of form-focused instruction on L2 pronunciation development: The role of explicit phonetic information. *Studies in Second Language Acquisition, 35* (1).

Saito, K. (in press, 2013b). Recasts in instructed L2 speech learning: Teaching the perception and production of English /r/ to adult Japanese learners. *Language Learning, 63*

Saito, K., & Brajot, F. (in press, 2013). Scrutinizing the role of length of residence and age of acquisition in the interlanguage pronunciation development of English /r/ by late Japanese bilinguals. *Bilingualism: Language and Cognition, 16*.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning, 62*, 595-633.

Schmitt, N. (2008). State of the art: Instructed second language vocabulary acquisition. *Language Teaching Research, 12,* 329-363.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research, 10*, 361-392.

Sheen, Y. (2007). The effects of corrective feedback, language aptitude and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301-322). Oxford: Oxford University Press.

Sheldon, A., & Strange, W. (1982). The acquisition of /ɹ/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3*, 243–261.

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 29,* 73-87.

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching, 44*, 225-236.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60,* 263-308.

Thomson, R. (2012). Improving L2 listeners' perception of English vowels: A computer–mediated approach. *Language Learning, 62*, 1231–1258.

Trofimovich, P., Collins, L., Cardoso, W., White, J., & Horst, M. (2012). A frequency-based approach to L2 phonological learning: Teacher input and student output in an intensive ESL context. *TESOL Quarterly*, *46*, 176-186.

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.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 56,* 431-462.

VanPatten, B. (2004). *Input processing and grammar instruction in second language acquisition*. Norwood N. J.: Ablex.

VanPatten, B., & Cardierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*, 225-243.

Walley, A. (2007). Speech learning, lexical reorganization, and the development of word recognition by native and non-native English speakers. In O. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 315–330). Amsterdam: John Benjamins.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*, 245-261.

Table 1. *Participant information by group*

	FFI+CF (n = 18)	FFI-only (n = 17)	Control (n = 14)
Age	M = 31.5 (SD = 10.6)	M = 31.0 (SD = 11.5)	M = 28.7 (SD = 8.0)
Gender	3 males, 15 females	5 males, 12 females	1 male, 13 females
LOR in months	M = 3.4 (SD = 8.8)	M = 7.6 (SD = 23.2)	M = 5.6 (SD = 9.9)
$x = 0$	12 learners	13 learners	8 learners
$0 < x \leq 12$	5 learners	3 learners	5 learners
$12 < x \leq 36$	1 learner	0 learner	1 learner
$36 < x$	0 learner	1 learner[a]	0 learner

Note. LOR, length of residence; [a]This learner reported 8 years of LOR

Table 2. *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

* indicates the 16 words included in the pre/post tests

Table 3. *50 tokens in the perception tests in relation to following vowel conditions*

		Front vowels	Central vowels	Back vowels
High vowels	(Trained items)	"rink, link" "reef, leaf" "read, lead"		"roof, Loof" "rule, lure" "room, loom"
	(Untrained items)	"ring, ling" "reach, leach"		"rude, lude" "root, loot"
Mid vowels	(Trained items)	"race, ace" "rent, lent" "rain, lane"		"road, load" "wrong, long" "roan, loan"
	(Untrained items)	"rate, late" "red, led"	"rough, laugh" "rush, lush"	"roll, loll" "rope, lope"
Low vowels	(Trained items)		"right, light" "ride, lied" "rice, lice"	
	(Untrained items)			

Table 4. *25 tokens in the controlled and spontaneous production tests in relation to following vowel conditions*

A. Controlled production test		Front vowels	Central vowels	Back vowels
High vowels	(Trained items)	rink, reef		rule, room
	(Untrained items)	reach		rude
Mid vowels	(Trained items)	race, rent		road[a], wrong
	(Untrained items)	rate	rough	roll
Low vowels	(Trained items)		Ryan, right	
	(Untrained items)			

B. Spontaneous production test		Front vowels	Central vowels	Back vowels
High vowels	(Trained items)	read		roof
	(Untrained items)	ring		route
Mid vowels	(Trained items)	rain		road[a]
	(Untrained items)	red	rush	rope
Low vowels	(Trained items)		rice	
	(Untrained items)			

Note. [a] "Road" was tested twice both in the controlled and spontaneous production tests.

Table 5. *Summary of acoustic values at pre-tests*

		Controlled Production				Spontaneous Production			
		Trained items		Untrained items		Trained items		Untrained items	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
FFI+CF (*n* = 18)	F3 (Bark)	14.75	0.69	14.70	0.75	14.88	0.64	14.77	0.58
	F2 (Bark)	11.74	0.92	11.58	0.97	11.90	0.96	11.90	0.96
	F1 (Bark)	4.63	0.95	4.01	0.55	4.55	0.63	4.55	0.63
	Transition duration (ms)	28.30	16.08	29.22	22.02	28.69	16.80	23.27	12.01
FFI-only (*n* = 17)	F3 (Bark)	14.36	0.69	14.42	0.78	14.54	0.66	14.60	0.80
	F2 (Bark)	11.10	1.00	11.28	1.18	11.21	0.92	11.29	0.86
	F1 (Bark)	4.26	0.62	4.19	0.54	4.52	0.64	4.31	0.56
	Transition duration (ms)	34.42	14.71	29.30	15.31	33.86	15.08	30.08	15.96
Control (*n* = 14)	F3 (Bark)	14.74	0.91	14.72	0.95	14.64	0.95	14.61	0.87
	F2 (Bark)	11.73	1.27	11.69	1.30	11.57	1.17	11.40	1.13
	F1 (Bark)	4.50	0.48	4.25	0.59	4.50	0.60	4.45	0.53
	Transition duration (ms)	30.38	20.00	28.59	20.19	32.47	21.84	28.01	18.16
NS Baseline (*n* = 10)	F3 (Bark)	12.06	0.15	12.06	0.27	12.24	0.40	12.29	0.32
	F2 (Bark)	10.04	0.38	10.09	0.45	10.33	0.49	10.40	0.53
	F1 (Bark)	4.31	0.37	4.10	0.39	4.35	0.28	4.35	0.28
	Transition duration (ms)	87.38	16.10	90.36	18.16	66.00	13.82	66.80	12.48

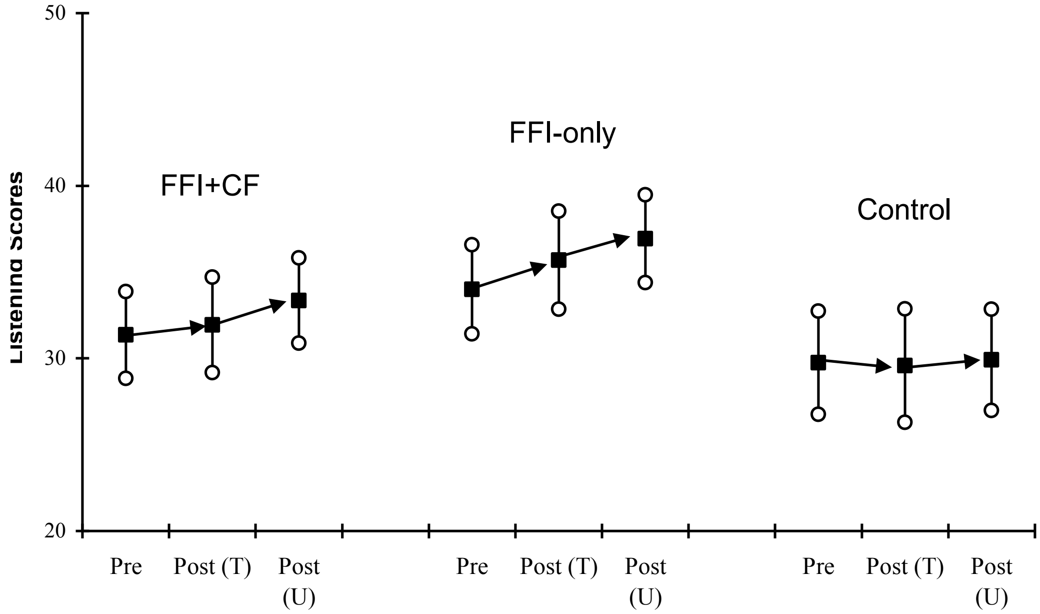

Figure 1. 95% confidence intervals and mean values of the learners' pre-post test scores. Post (T) denotes their performance of materials spoken by their instructors. Post (U) denotes their performance of materials spoken by untrained talkers.

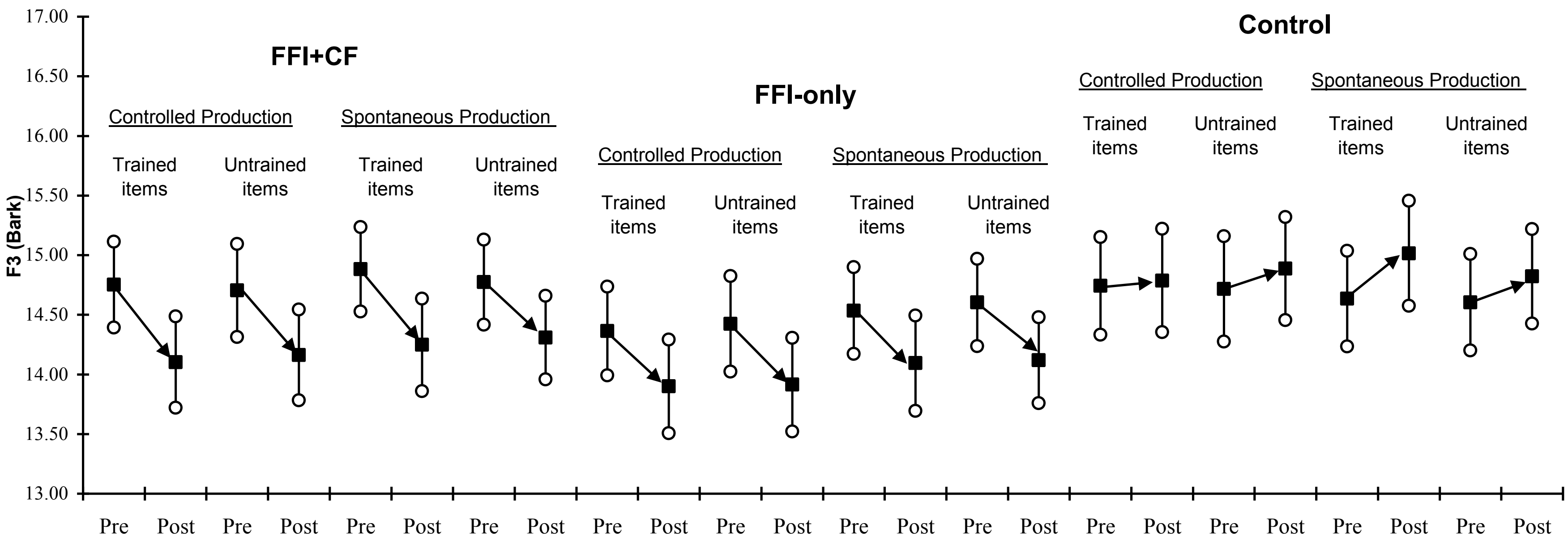

Figure 2. 95% confidence intervals and mean values of F3 (Bark) in Group × Task × Lexical contexts. Both of the FFI+CF and FFI-only groups outperformed the control group at all contexts.

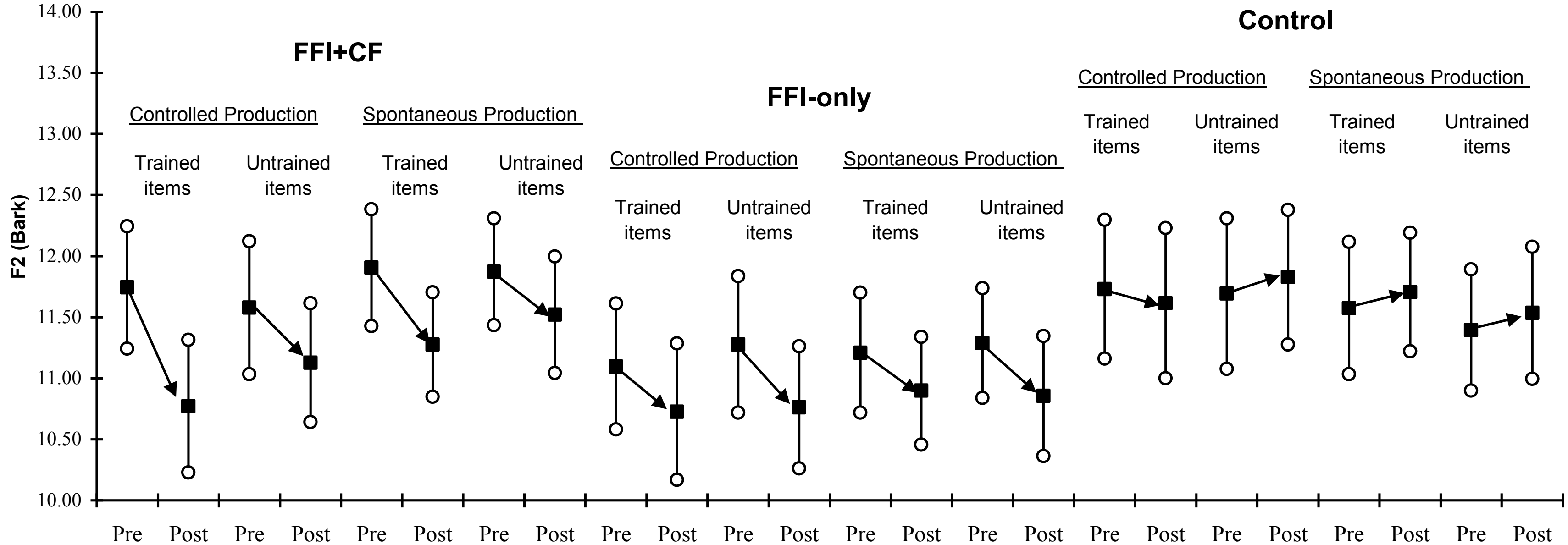

Figure 3. 95% confidence intervals and mean values of F2 (Bark) in Group × Task × Lexical contexts. Both of the FFI+CF and FFI-only groups outperformed the control group at all contexts.

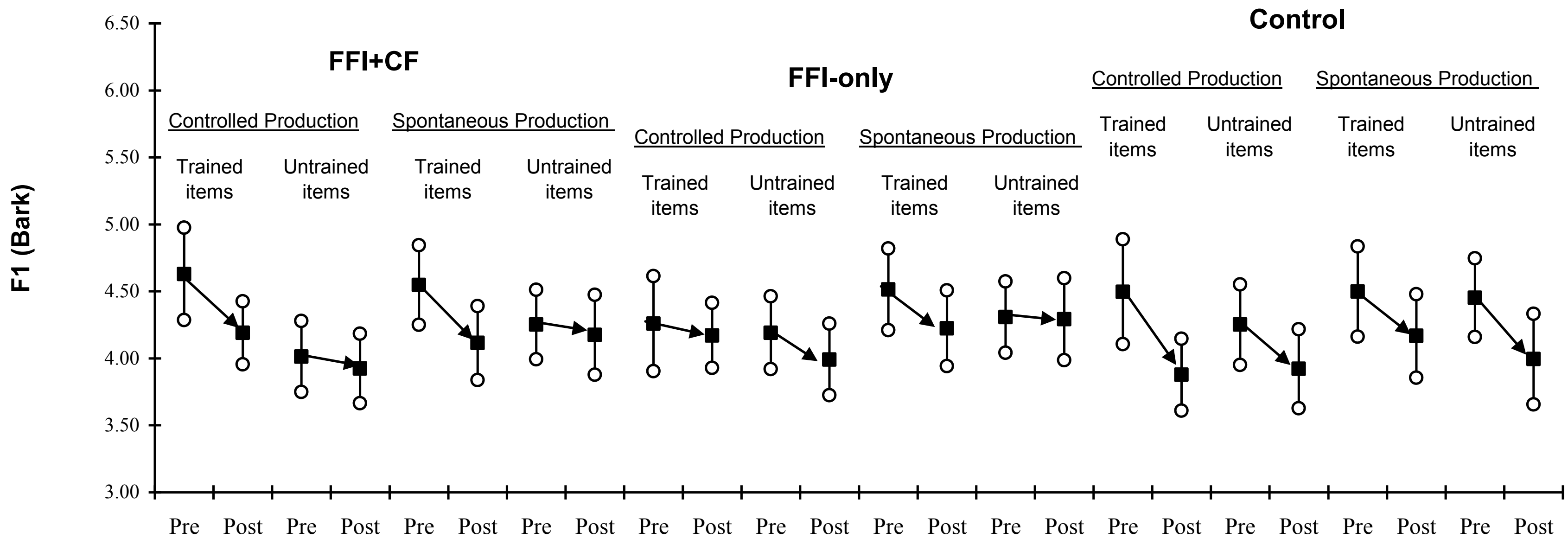

Figure 4. 95% confidence intervals and mean values of F1 (Bark) in Group × Task × Lexical contexts. All groups equally lowered F1 values after instruction.

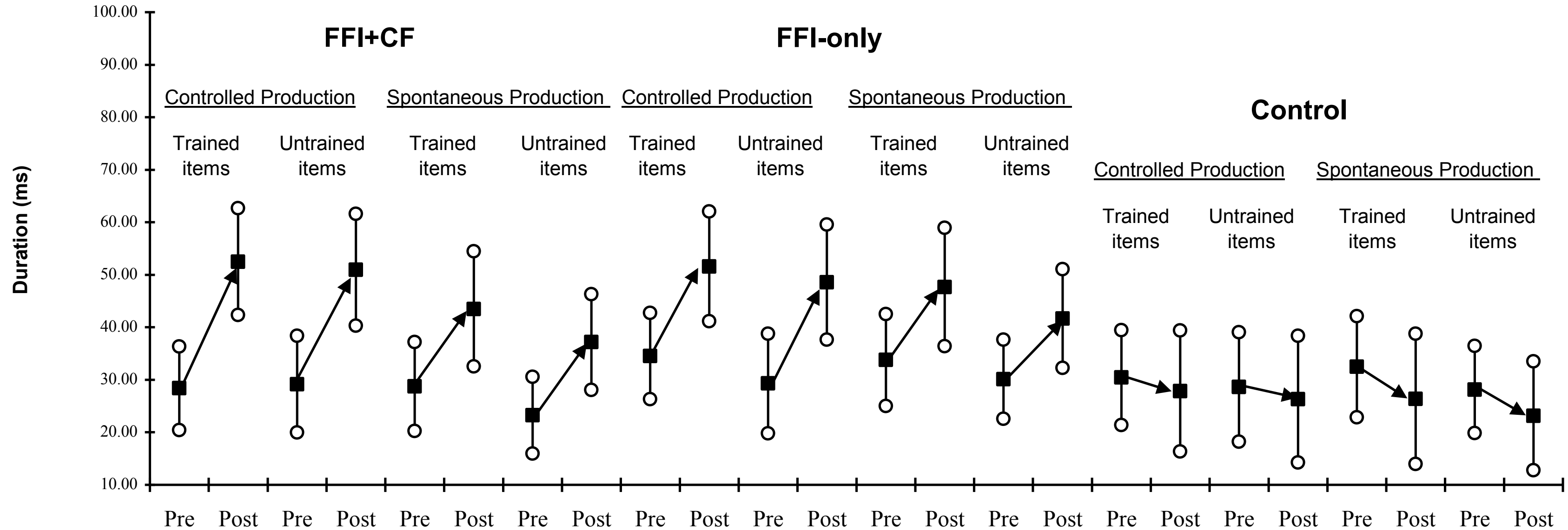

Figure 5. 95% confidence intervals and mean values of transition duration (ms) in Group × Task × Lexical contexts. Both of the FFI+CF and FFI-only groups outperformed the control group at all contexts.

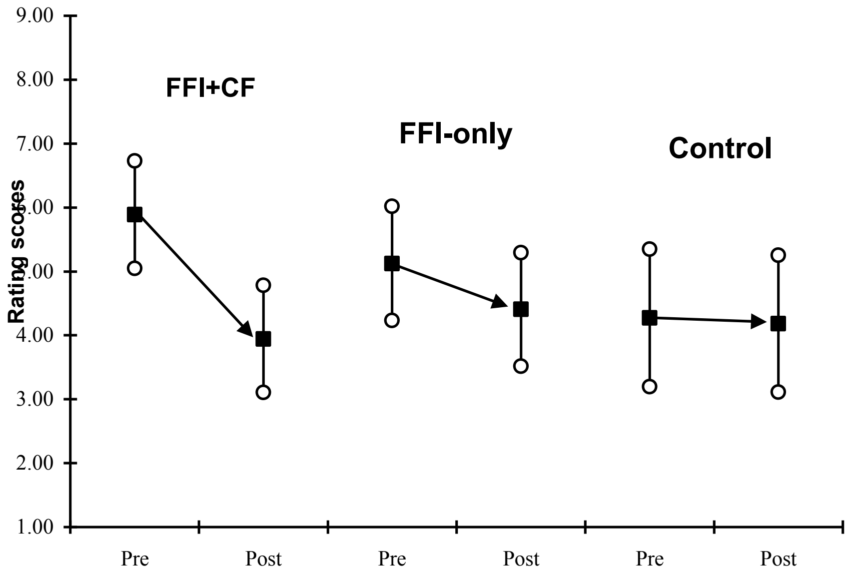

Figure 6. 95% confidence intervals and mean values of teacher evaluation in Group × Task contexts. Both of the FFI+CF and FFI-only groups showed a significant improvement over time.