

# REEXAMINING EFFECTS OF FORM-FOCUSED INSTRUCTION ON L2 PRONUNCIATION DEVELOPMENT

## *The Role of Explicit Phonetic Information*

Kazuya Saito

*Waseda University*

---

The present study examines whether and to what degree providing explicit phonetic information (EI) at the beginning of form-focused instruction (FFI) on second language pronunciation can enhance the generalizability and magnitude of FFI effectiveness by increasing learners' ability to notice a new phone. Participants were 49 Japanese learners of English in English as a foreign language setting. Whereas the control group ( $n = 14$ ) received meaning-oriented lessons without any focus on form, the experimental groups received 4 hr of FFI treatment designed to encourage them to practice the target feature of an English /ɪ/ in meaningful discourse. Instructors provided EI (i.e., multiple exposure to an exaggerated model pronunciation of /ɪ/ and rule presentation on the relevant articulatory configurations) to the FFI+EI group ( $n = 17$ ) but not to the FFI-only group ( $n = 18$ ). Their

The results reported here are based on my doctoral dissertation submitted to McGill University in 2011. I would like to express my deepest gratitude toward Roy Lyster for his advice and guidance as my dissertation supervisor at each phase of the entire project. I also wish to thank Murray Munro, Pavel Trofimovich, Leila Ranta, Walcir Cardoso, and the four anonymous *SSLA* reviewers for their helpful comments on the earlier versions of the manuscript. Finally, I am grateful to the two teachers, Yukio Nagato and Anna-Marie Staniewski, and all the volunteer students for their participation in the project as well as to Noriaki Okamura at ECC Foreign Language Institute, for his accommodation in coordinating and providing me with access to classrooms and relevant facilities at the school.

Correspondence concerning this article should be addressed to Kazuya Saito, School of Commerce, Waseda University, 1-6-1 Nishi Waseda Shinjuku, Tokyo 169-8050, Japan; e-mail: kazuya.saito@waseda.jp.

pre- and posttest performance was acoustically analyzed according to various lexical, task, and following vowel conditions. The results of the ANOVAs showed that (a) the FFI-only group demonstrated moderate improvement with medium effects (e.g., change from hybrid exemplars to poor exemplars), particularly in familiar lexical contexts, and (b) the FFI+EI group not only demonstrated considerable improvement with large effects (e.g., change from hybrid exemplars to good exemplars) but also generalized the instructional gains to unfamiliar lexical contexts beyond the instructional materials.

---

Form-focused instruction (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). It is hypothesized to be most effective when integrated into communicative-oriented and content-based classrooms because second language (L2) learners can notice and practice target linguistic features during meaningful discourse, which in turn enhances (a) their form-meaning mappings (Doughty, 2003; Ellis, 2002; VanPatten, 2004) and (b) their gradual transition from effortful to automatic use of rules (DeKeyser, 2003, 2007; Lyster, 2007). In particular, Lyster and Ranta developed a pedagogical sequence of FFI in relation to three stages of interlanguage development: noticing → awareness → practice (Lyster, 2007; Ranta & Lyster, 2007). According to this sequence, learners should first be given some preplanned focused tasks that are designed to promote their noticing of the target feature in L2 input, especially at the initial stage of interlanguage development—namely, the noticing phase (VanPatten, 2004). Second language learners subsequently need to develop more targetlike representations, especially through both input- and output-based activities, for the purpose of deeper and more elaborate processing of form—namely, the awareness phase (Swain, 2005). Finally, after learners successfully restructure and develop interlanguage representations, they are ready to engage in FFI activities to repetitively practice the target feature in production under communicatively authentic contexts to proceduralize their declarative knowledge—namely, the practice phase (DeKeyser, 2003, 2007).

Second language acquisition research into FFI, however, has primarily targeted morphosyntactic development and has left a huge gap in research on the effects of instruction on pronunciation (see Derwing & Munro, 2005, for discussion). Saito and Lyster (2012) took a first step toward testing how a range of FFI techniques can promote the acquisition of the English sound /ɪ/ by adult Japanese learners. Although the results showed that L2 pronunciation development can be amenable to FFI, they also revealed several limitations (i.e., moderate improvement only within familiar lexical items). To this end, the current study was

designed to examine whether and to what degree providing explicit phonetic information (EI) at the beginning of FFI lessons can enhance the generalizability and magnitude of the effects of FFI. In the following sections, I first review how L2 learners detect new sounds from L2 input and develop new phonetic categories in naturalistic settings. I then describe to what degree FFI can facilitate the default L2 speech learning process by summarizing the findings as well as limitations of the Saito and Lyster study. Finally, I discuss the possibility of adding EI to the beginning of FFI lessons as a remedial technique.

## **RELATED LITERATURE**

### **L2 Speech Learning**

Whereas L2 speech studies have extensively examined how L2 learners create new phonetic categories in their long-term memory representation in relation to the quality and quantity of L2 input, a great deal of attention has also been given to investigating how the nature of the category changes over time according to different stages of L2 development (e.g., perceptual assimilation model: Best & Tyler, 2007; Bundgaard-Nielsen, Best, & Tyler, 2011a, 2011b; speech learning model: Flege, 1995, 2003, 2009; lexical restructuring model: Walley, 2007). At the initial stage of L2 speech learning, L2 learners pay primary attention to prosodic patterns of language and start recognizing words to derive meaning from ambient aural input (Cutler, Dahan, & van Donselaar, 1997; Kuhl, 2000, 2004). They decode word-sized units of L2 phonological information, and their perception and production of L2 sounds tends to be influenced by lexical factors such as familiarity, frequency, and density (Bradlow & Pisoni, 1999; Imai, Walley, & Flege, 2005).

As their vocabulary size increases, however, these learners will be forced to attend to fine-grained phoneme discrimination and identification to perceive and produce phonetically similar words such as minimal pairs. Although they are sensitive primarily to word-sized units of L2 phonological information, they concurrently become more aware of sound-sized units of L2 phonological information (i.e., segmental aspects of the speech stream). At this stage, such L2 learners with large vocabulary size can show highly consistent patterns of sound recognition (i.e., assimilation and dissimilation of L2 sounds relative to L1 system: Best & Tyler, 2007; Bundgaard-Nielsen et al., 2011a, 2011b), and their segmental performance tends to be less subject to the influence of lexical factors (Flege, Frieda, Walley, & Randazza, 1998; Walley, 2007). Flege (1995, 2003, 2009) claimed that this phonetic-level noticing of perceptual aspects of new segmental sounds crucially leads learners to (a) create new phonetic categories in long-term memory and (b) generalize the

newly acquired phonetic knowledge from familiar to new lexical contexts (see also Kuhl, 2000, 2004).

Finally, when the category is integrated into a learner's developing L2 system, many researchers—including Flege—claim that change occurs first in the perception domains, which, in turn, activates relevant sensorimotor skills for production abilities (for details of the perception-first view, see Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Zhang et al., 2009). In conjunction with the default L2 speech learning process, pronunciation instruction should therefore be designed to (a) raise learners' perceptual noticing and awareness of new sounds not only at a lexical level but also at a phonetic level to promote the formation of new phonetic categories and (b) encourage them to practice the new L2 sounds in output to enhance their production abilities at various processing levels in relation to the present state of their developing mental representation.

### English /ɹ/

The current study highlights one of the most well-researched cases of L2 speech learning: the acquisition of English /ɹ/ by adult Japanese learners (for review, see Bradlow, 2008). Because the Japanese phonetic system has no counterpart approximants, Japanese learners tend to perceive English /ɹ/ and /l/ as poor exemplars of the Japanese tap, which is hypothesized to be in “a position in a phonological space that is somewhere between English /ɹ/, /l/, and /d/” (Flege, Takagi, & Mann, 1995, p. 25; see also Hattori & Iverson, 2009). Due to such complex phonetic contexts, Japanese learners continue to demonstrate a tremendous amount of resistance toward acquiring the English /ɹ/-/l/ contrast both in perception (Lively, Logan, & Pisoni, 1993) and in production (Larson-Hall, 2006).

From an acoustic standpoint, although native speaker (NS) listeners draw on variance in third formant (F3) frequencies to perceive the /ɹ/-/l/ contrast (i.e., English /ɹ/ < 2,000 Hz, and English /l/ > 2,400 Hz; see Espy-Wilson, 1992; Espy-Wilson, Boyce, Jackson, Narayanan, & Alwan, 2000), Iverson and his colleagues have found that Japanese learners ignore F3 values and instead use variance in second formant (F2) frequencies as their default strategy (see Iverson et al., 2003). Furthermore, Bradlow (2008) argued that their insensitivity to F3 values hinders activation of new articulatory configurations highly relevant for producing English /ɹ/ (i.e., the combination of lip rounding and simultaneous constrictions in the pharyngeal and glottal area of the vocal tract), which in turn leads to their nontargetlike pronunciation (for empirical evidence, see Lotto, Sato, & Diehl, 2004). In regard to this specific L2

learning instance, the current study further explores to what degree meaning-oriented pronunciation instruction—FFI with and without EI—can help learners create new phonetic categories and produce new sounds in a range of lexical, phonetic, and task contexts in an efficient and effective manner.

## **FFI and L2 Pronunciation Development**

With a general absence of research that specifically investigates the pedagogical capabilities of FFI in L2 pronunciation development, Saito and Lyster (2012) examined how FFI with and without corrective feedback (CF) can be facilitative of L2 pronunciation development of /ɪ/ by Japanese learners. The results showed that (a) exposing learners to form-focused tasks alone was not sufficient to promote acquisition and (b) providing CF (i.e., recasts) to learners' mispronunciation during communicative tasks plays an important role in leading to change in their L2 pronunciation performance. The critical role of CF in FFI effectiveness was attributed to the dual pedagogical function of pronunciation-focused CF. Second language learners were pushed to practice correct pronunciation forms (i.e., pronunciation practice) while carefully listening to a teacher's model pronunciation of English /ɪ/ (i.e., listening practice). However, the study also generated several questions and was constrained by certain limitations worthy of further research.

First, all of the relevant findings should be interpreted with caution because it still remains unclear whether and to what degree the learners could generalize changes resulting from FFI to unfamiliar lexical tokens. Learners' gains were found only when their performance was tested via familiar items, which they practiced during 4 hr of FFI treatment. This point could be used as evidence that the learners who received CF succeeded in restructuring mental representations at a lexical level (i.e., lexically driven L2 phonological development) but failed to do so at a phonetic level (i.e., phonetically driven L2 phonological development), which is hypothesized to be necessary for the development and generalization of new phonetic categories to unfamiliar lexical contexts (Flege, 1995, 2003, 2009). At the same time, however, the lack of statistical significance in the generalizability task could be attributed to the construct validity of its design; that is, generalizability was checked simply by asking learners to read a list of four words only at the posttest sessions. In this respect, in the current study learners were asked to read a number of familiar and unfamiliar tokens via controlled and spontaneous production tests at both pre- and posttest sessions. Finally, the amount of improvement resulting from FFI in the Saito and Lyster (2012) study could be considered moderate rather than large. A close

examination of the acoustic analyses showed that even the learners receiving CF changed their mean F3 values only within the range of the hybrid-poor exemplars of English /ɪ/ (e.g., 2,400–2,500 Hz, 14–15 Bark). To this end, the current study tests the pedagogical efficacy of providing EI at the beginning of FFI lessons to enhance the generalizability and magnitude of the effects of FFI on both familiar and unfamiliar lexical items across various task and phonetic conditions.

### **Adding EI to FFI**

In spite of a great deal of research on the effects of EI on L2 morphosyntactic development (Norris & Ortega, 2000; Spada & Tomita, 2010), Henry, Culman, and VanPatten (2009) claimed that it still remains unclear to what extent EI—as an independent variable—impacts L2 acquisition processes, arguably because the role of EI has been “conflated with the issue of explicit versus implicit learning” (p. 561) and has been tested in tandem with decontextualized drill activities in primary studies. Although few in number, recent SLA studies have begun to tease apart and test the role of EI in more meaning-oriented FFI lessons in which learners are to process language for meaning so that their explicit understanding of problematic structures enhances the subsequent acquisition of implicit knowledge (Fernandez, 2008; Henry et al., 2009; for further discussion, see Ellis, 2002). Spada and Lightbown (2008) speculated that EI should be given before (or after) meaning-oriented lessons when the target language features are almost unnoticeable, especially during ordinary communicative interaction, “either because they are acoustically imperceptible (e.g., most grammatical morphology in English) or redundant and unlikely to affect comprehension (e.g., word order in English questions)” (p. 186).

Derwing and Munro (2005) emphasized the importance of EI in the context of pronunciation teaching and claimed that “students learning L2 pronunciation benefit from being explicitly taught phonological form to help them notice the difference between their own productions and those of proficient speakers in the L2 community” (p. 388). When used to teach new segmental sounds, EI consists of multiple exposures to a teacher’s model pronunciation of the target sounds followed by an explanation of relevant articulatory configurations to “raise learners’ consciousness” (Celce-Murcia, Brinton, & Goodwin, 1996, p. 36). Given that pedagogical options for pronunciation teaching have been exclusively limited to audio-lingual teaching techniques such as minimal pair drills and repetition activities (for discussion, see Trofimovich & Gatbonton, 2006), the role of EI has been tested only in relation to these decontextualized practice opportunities. The results of previous studies have generally revealed that the combination of EI and decontextualized

instructional approaches impacts learners' pronunciation development only at a controlled-speech level but not at a spontaneous-speech level (e.g., Elliott, 1997). However, little research attention has been given to the relationship between EI and FFI in the context of pronunciation teaching; the current study takes a first step to examine this relationship.

It is reasonable to assume that EI could be an important variable for L2 phonological development, especially in conjunction with more meaning-oriented FFI. Given that L2 learners generally process L2 phonological information at a lexical level, providing EI at the beginning of FFI lessons will immediately draw learners' attention to sound-sized units of L2 phonological information. In this way, EI is hypothesized to promote their noticing of the perceptual difference between a new L2 sound and its L1 counterpart which, many researchers argue, could be a first step toward L2 phonological development (Flege, 1995, 2003, 2009; Kuhl, 2000, 2004). In a similar manner, teaching articulatory gestures with listening discrimination activities could also help L2 learners notice the perceptual aspects of the new L2 sound (Celce-Murcia et al., 1996). This phonetic-level restructuring ultimately leads L2 learners to establish and internalize new phonetic categories as well as to generalize the newly acquired phonetic knowledge from familiar to new lexical contexts.

Another factor affecting EI is the differential learnability of L2 sounds; that is, L2 learners can quickly create new phonetic categories even without much modified input according to the acoustic characteristics of new L2 sounds (e.g., temporal vs. spectral differences). Explicit phonetic information may even be unnecessary for these relatively easy features (see Flege, 1989, for the impacts of intensive perceptual training on the acquisition of the final stop deletion by Chinese learners of English). Some L2 sounds, however, are extremely difficult and time consuming to acquire, such as the nonnative English /ɪ/–/l/ contrast for Japanese learners (Bradlow, 2008). In short, it is hypothesized that EI is highly beneficial for L2 pronunciation development of /ɪ/ by L1 Japanese learners.

## THE STUDY

The primary goal of the current study is to examine to what degree adding EI (i.e., triggering phonetically driven L2 phonological learning) to FFI (i.e., triggering lexically driven L2 phonological learning) can enhance the generalizability and magnitude of the overall instructional impact on interlanguage development of /ɪ/ by Japanese learners. Following the original definition (Spada, 1997), FFI in the current study included a broad range of instructional options (i.e., focused tasks and CF), which can be either explicit or implicit, or proactive or reactive. However, EI is defined from a L2 phonology perspective as a separate explicit intervention to draw learners' attention from lexical units as a whole toward segmental



aspects of L2 speech (see also Spada & Lightbown, 2008, for isolated FFI in a grammar-teaching context). Thus, this study is guided by two principal research questions:

1. To what degree is FFI facilitative of L2 pronunciation development of /ɪ/ by Japanese learners of English in both familiar and unfamiliar lexical contexts?
2. To what degree is FFI+EI facilitative of L2 pronunciation development of /ɪ/ by Japanese learners of English in both familiar and unfamiliar lexical contexts?

## METHOD

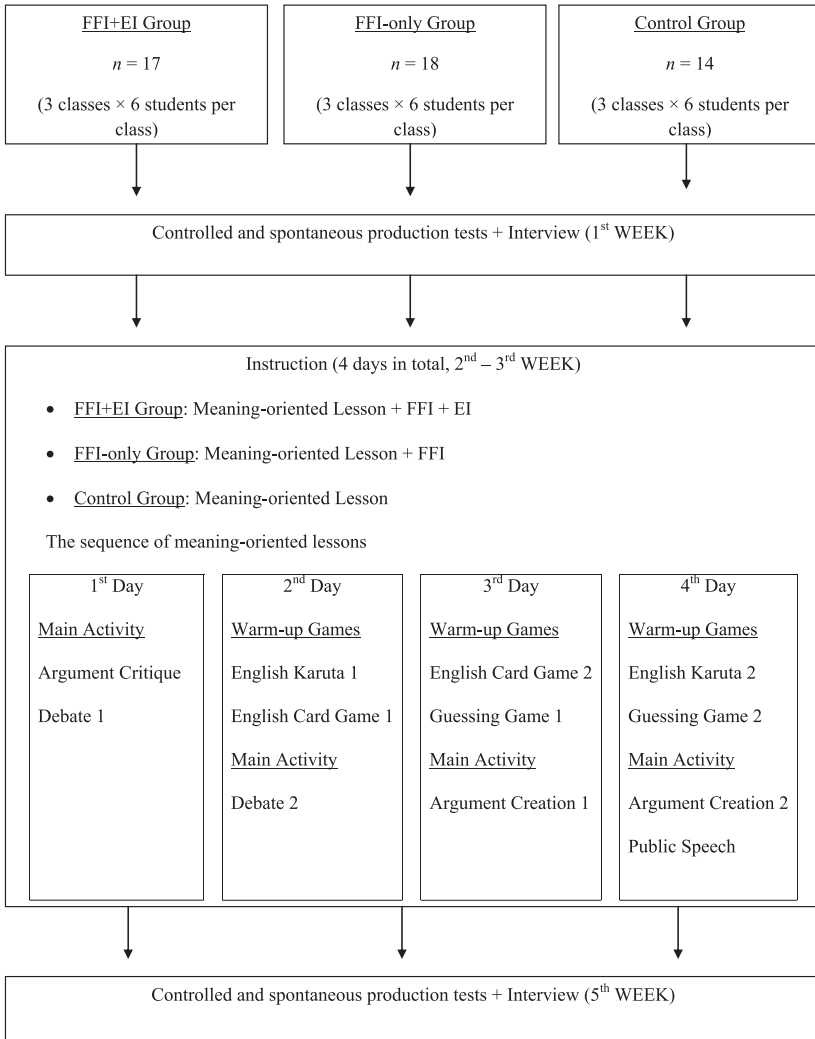
### Design

Participants in the experimental groups received a range of form-focused tasks embedded in 4 hr of meaning-oriented lessons to encourage learners to notice and practice the target sound feature (i.e., English /ɪ/) during meaningful discourse. During these activities, the instructors also provided CF following students' mispronunciation or unclear pronunciation of /ɪ/. Only students in the FFI+EI group received EI at the beginning of FFI lessons. For the control group, students received meaning-oriented lessons that were comparable in terms of duration and content but without any focus on form (i.e., English /ɪ/). Instructional treatments consisted of four 1-hr lessons distributed over 2 weeks (1-hr lesson × 2 times per week × 2 weeks = 4 hr). All classes were videotaped and observed by the researcher, who always sat at the back of the classroom to ensure the consistency of FFI treatment for the entire project. Two weeks after the end of the lessons, all students took posttests and were interviewed.<sup>1</sup> Figure 1 summarizes the design of the study and the procedures followed.

### Participants

**Students.** The project was conducted at a private language institute in Osaka, Japan. For the purpose of student recruitment, the author created ads that advertised free 4-hr English lessons with a focus on developing English argumentative skills. The ads were posted online on a number of English education web sites and a few social network web sites. The private language institute also distributed the ads to all of their students. Interested participants contacted the author through e-mail or by phone to set up a date for their first interview and pretest sessions. Even though 54 students—18 students per group—initially participated in the current project, 5 did not complete the instructional





**Figure 1.** Summary of the procedure.

treatment nor attend the posttest sessions for several personal reasons. A total of 49 participants (males,  $n = 7$ ; females,  $n = 42$ ) were included in the final analysis (age:  $M = 29.04$ ,  $SD = 8.64$ ).

According to the first interview, although some participants reported having lived in English-speaking countries for a few years, most of the participants had never been abroad (length of residence [LOR]:  $M = 4.44$  months,  $SD = 8.66$ ). Even though all of the learners had received 6–10 years of English education at public schools in Japan, they stated that they had no opportunity to actually speak English in communication at the time of

the project, except for a few hours of English lessons either at the language institute or their college-level schools. In this respect, unlike the Saito and Lyster (2012) study, which involved intermediate Japanese learners with a mean LOR exceeding 1 year in English as a second language (ESL) settings (LOR:  $M = 15.5$  months,  $SD = 31.8$  months), the current study, because it involved beginner-intermediate Japanese learners in English as a foreign language (EFL) settings, could isolate and measure the pure impact of FFI on their phonological development at the initial stage. The 49 participants were randomly assigned to nine classes of 6 students. Two treatment groups and one control group, each of which comprised three classes, were formed: (a) FFI+EI group (three classes,  $n = 17$ : males,  $n = 3$ ; females,  $n = 14$ ), (b) FFI-only group (three classes,  $n = 18$ : males,  $n = 3$ ; females,  $n = 15$ ), and (c) control group (three classes,  $n = 14$ : male,  $n = 1$ ; females,  $n = 13$ ).

**Instructors.** Two experienced NSs of English (i.e., one male from California and one female from Ontario, Canada), who were employed as teachers by the language institute, participated in the study. Both teachers were selected by the language institute on the basis of their extensive EFL teaching experience in Japan. One instructor taught five classes (i.e., two FFI-only classes and three control classes), and the other taught the other four (i.e., one FFI-only class and three FFI+EI classes).<sup>2</sup>

## Interview

All participants were individually interviewed right after the pretest sessions (i.e., the initial interview) as well as the posttest sessions (i.e., the final interview). The purpose of these interviews was to elicit personal information including age and English learning experiences at the onset of the project as well as to ascertain the degree to which learners had focused on form and meaning during the FFI lessons in a retrospective manner after the project was completed. All communication was in Japanese and audiorecorded.

## FFI Treatment

The FFI treatment used in the current study has been described in detail in the Saito and Lyster (2012) study. In this section, therefore, only a brief description of training methods is provided.

**Target Words.** In the current study, the instructional treatment highlights 39 target words that include /ɹ/ in various phonetic contexts:

**Table 1.** Thirty-nine target words included in the instructional treatment

Phonetic contexts	Target words
Word-initial	* <i>race</i> , * <i>rain</i> , <i>ram</i> , <i>rat</i> , * <i>read</i> , * <i>reef</i> , * <i>rent</i> , * <i>right</i> , * <i>rice</i> , <i>ride</i> , * <i>rink</i> , <i>river</i> , * <i>road</i> , <i>roan</i> , <i>robot</i> , <i>rock</i> , <i>rocket</i> , <i>Rome</i> , * <i>roof</i> , * <i>room</i> , <i>round</i> , * <i>rule</i> , <i>run</i> , * <i>Ryan</i> , <i>wrap</i> , * <i>wrong</i>
Word-medial	<i>arrive</i> , <i>correct</i> , <i>pirate</i>
Consonant cluster	<i>bread</i> , <i>crab</i> , <i>crime</i> , <i>crowds</i> , <i>free</i> , <i>fries</i> , <i>fruit</i> , <i>grass</i> , <i>green</i> , <i>pray</i>

Note. \* = the 14 words included in the pre- and posttests.

26 words for word-initial positions, 3 words for word-medial positions, and 10 words for consonant clusters. It is important to note that all of these 39 words were minimally paired—or nearly minimally paired—with /l/, which required the students to make a clear distinction between the /ɹ/-/l/ contrast for the purpose of their successful L2 communication (see Table 1).

**Instructional Treatment.** The primary purpose of the meaning-oriented lessons was to acquire English argumentative skills that entailed logical thinking, negotiation and debating skills, and public-speaking abilities. All instructional materials were color printed and labeled as one package for each lesson for each pair of learners. To successfully complete each activity, learners had to use a few lexical items, including /ɹ/ in various phonetic positions, and all of these words were highlighted in red to draw learners’ attention to the target feature of /ɹ/ (i.e., typographically enhanced input, see Han, Park, & Combs, 2008). For example, under the topic *English debating*, learners were divided into two groups, and each team supported or objected to given topics (e.g., *Is **running** inside better than **running** outside? Is it good to have a **rainy** day?*). As warm-ups for the main activities, the learners also played several communicative games. For example, in English Karuta, which is based on a traditional Japanese card game, 36 cards were placed on the table. Each card represented one lexical item and portrayed a relevant picture along with the first letter of the word. These 36 lexical items were minimally paired with English /ɹ/-/l/ contrast (e.g., *reef* vs. *leaf*). When the instructors read a list of these words, the learners had to pick up the card as soon as possible.

**Pronunciation-Focused Recasts.** Corrective feedback treatment was operationalized as a form of pronunciation-focused recasts; that is, the instructors reformulated learners’ mispronunciation or unclear

pronunciation of /ɪ/ without altering the original meaning of the utterance. To increase the saliency of the corrective message in recasts, the instructors were asked to consistently recast only one word using falling intonations. Such recast techniques were labeled as partial recasts (Sheen, 2006) as shown in (1).

(1) Argument critique:

Student: *I love to eat rice* /laɪs/\*.

Teacher: *Rice* /ɪaɪs/.

Student: *Rice* /ɪaɪs/. *I have tried many kinds of rice* /ɪaɪs/.

## El Treatment

For the FFI+EI group, for 5 of the first 10 min of Day 1 and the first 5 min of Days 2, 3, and 4, the instructor started with EI on how to perceive and produce /ɪ/. Given that FFI treatment is hypothesized to promote lexically driven L2 phonological development, the primary purpose of EI was to induce learners' awareness and noticing of the perceptual difference between /ɪ/ and /l/ at a phonetic level. Because /l/ is acoustically similar to the Japanese tap /ɾ/ (e.g., Hattori & Iverson, 2009), the instructor provided his exaggerated pronunciation model of the /ɪ/-/l/ contrast, highlighting its perceptual difference (i.e., acoustic enhancement).<sup>3</sup> At the same time, the instructor taught relevant articulatory configurations about /ɪ/, which enabled learners to actually produce the new sound and encouraged them to listen to perceptual aspects of their own production to help them notice the acoustic—and articulatory—difference between English /ɪ/ and Japanese tap /ɾ/ (see Diehl, Lotto, & Holt, 2004, for the motor theory of speech perception).<sup>4</sup> The specific procedures of the EI treatment on Day 1 were as follows:

1. The instructor asked learners to carefully listen to his exaggerated pronunciation of both /ɪ/ and /l/ at a phonetic level so that they would become aware first of the perceptual characteristics of /ɪ/ and then of its perceptual difference relative to /l/, which is acoustically similar to the Japanese tap /ɾ/.
2. Subsequently, the relevant articulatory configurations (i.e., lip rounding, the raising of the tongue tip, and obstruction in the glottal and pharyngeal areas of the vocal tract<sup>5</sup>) were explained for /ɪ—but not /l—with the aid of an articulatory diagram. In particular, the instructor emphasized the relative importance of lip rounding following Bradlow's (2008) recommendation<sup>6</sup> as well as recent L2 speech research findings (see also the ultrasound and optical tracking of lip constrictions for producing word-initial /ɪ/, Campbell, Gick, Wilson, & Vatikiotis-Bateson, 2010).

3. Finally, learners also performed three types of production practice: The learners first repeated only the instructor's model pronunciation of /ɪ/ (i.e., phonetic practice) and then produced and perceived minimally paired words (i.e., lexical practice) and sentences filled with target sounds (i.e., sentence practice).

For the rest of the lessons (Day 2–Day 4), this practice phase was eliminated due to time constraints. Instead, the instructor always gave very short metalinguistic explanations (i.e., around 5 min long) at the beginning of the class, by (a) reminding students of the relevant articulatory gestures for /ɪ/, (b) modeling it with exaggeration, and (c) asking them to carefully listen and repeat. To ensure that all groups received the same amount of instruction (i.e., 4 hr), the instructors were asked to spend more time on warm-up games (for the FFI-only group) and small talk (for the control group) instead of providing EI.

### Control Group

The 14 participants in the control group also received comparable meaning-oriented lessons on English argumentative skills but with neither FFI nor EI; the students received feedback not on any pronunciation errors but rather on ungrammatical or inappropriate lexical choices (e.g., *You should say, "I dropped a pen" instead of "I fell a pen"*) as well as the content of the lessons (e.g., *Your opinion could be more convincing if you touched upon the opponent's critique*). As for warm-up games, the participants in the control group were given different communicative games without any focus on pronunciation or listening practice, which the instructor usually used in her regular English conversation classes.

### Teacher Training

Two instructors participated in a total of 4 hr of teacher training led by the researcher over a 2-day period. First, the instructors were given (a) a set of guidelines that specified the objectives and procedures for all instructional materials and (b) a list of 39 target words textually highlighted in the FFI materials. Next, the researcher carefully explained (a) how to deliver each activity (i.e., English argumentative activities + communicative games) as well as (b) when and (c) how to provide CF to learners' mispronunciation of /ɪ/ (i.e., pronunciation-focused recasts). To help their understanding of the contexts, the researcher also demonstrated model lessons when necessary. Finally, the instructor that

taught all three of the FFI+EI classes also received special materials for EI and practiced demo lessons with the researcher.

## Measures

To measure the effects of the two types of FFI (i.e., FFI+EI vs. FFI-only) as compared to the control group, all learners were asked to complete two types of production tests at both pre- and posttest sessions: (a) the controlled production (CP) test (i.e., reading a list of words) and (b) the spontaneous production (SP) test (i.e., describing a set of pictures).

**Materials.** All target words were consonant-vowel-consonant (CVC) singletons except one word, *Ryan*, which was a CVVC (see the 14 words with asterisks in Table 1). Given that Japanese learners experience difficulties in producing and perceiving /ɹ/, particularly in word-initial positions (Bradlow et al., 1997; Sheldon & Strange, 1982), the underlying assumption for including word-initial tokens in the test materials was that examination of the relatively difficult phonetic contexts could reveal significant results, which can be generalized to other relatively easier contexts (e.g., postvocalic positions; for a similar discussion on perception training studies, see Lively et al., 1993). To measure the generalizability of FFI effectiveness, these CP and SP tests included (a) familiar items that the learners were exposed to during the FFI treatment and (b) unfamiliar items that never appeared in the FFI materials. Furthermore, with regard to familiar and unfamiliar items respectively, any instructional impact on the learners' performance of /ɹ/ was carefully examined in relation to two affecting variables: (a) task types (i.e., controlled- vs. spontaneous-speech levels) and (b) phonetic contexts (i.e., singletons with front or back vowels).

At both pre- and posttest sessions, all students first completed the SP test to measure their SP abilities without too much awareness of /ɹ/; they then moved on to the CP test. All of the pre- and posttest sessions with the 49 learners were individually administered in a quiet room. Their speech tokens were recorded by means of a Roland R05 Wave recorder, at a 44.1 kHz sampling rate and a 16-bit resolution. A Yamaha unidirectional DM-20SL microphone was used, and all of the recordings were stored on the hard drive of a laptop computer.

*SP test.* As a reliable outcome measure of SP abilities, Ellis (2002) recommended communicative free production, defined as "an activity that calls for unplanned language use directed at fulfilling some communicative purpose" (p. 225).<sup>7</sup> Timed picture-description tasks have been empirically validated as cognitively demanding (Derwing, Rossiter, Munro, & Thomson, 2004) and are commonly used in L2 morphosyntax

studies (e.g., Lyster, 2004). In this regard, the current study adopted timed picture-description tasks to measure SP to assess to what degree L2 learners can accurately produce new L2 sounds (i.e., /ɹ/) when their attention is focused on communicating a message. Participants were asked to describe 10 pictures with 6 distractor pictures, and each of the 10 pictures led the learners to pronounce one target word, including word-initial /ɹ/ (i.e., 5 familiar items + 5 unfamiliar items). In total, 980 tokens (i.e., 10 words × 49 students × 2 test sessions = 980 tokens) were produced at both the pre- and posttest sessions. These words are listed in Table 2.

The task was operationalized as follows:

1. Learners were first given 10 s to memorize four key words on a sheet of paper that related to two pictures they were to describe; one of the two key words for each picture was a target word including /ɹ/ in word-initial position, whereas the other was a distractor.
2. Right after the card was taken away, they were given one picture after another to describe and were required to use all of the key words they had just memorized. To minimize their use of explicit knowledge, the learners were prompted to complete the task without any planning time; that is, they described the pictures as soon as they received them.
3. After describing the pictures, they moved on to the next four key words for another set of two pictures.

*CP test.* To measure their L2 pronunciation performance of /ɹ/ at controlled-speech levels, learners were asked to read a list of 40 words in total, 15 of which were target tokens (i.e., 10 familiar items + 5 unfamiliar items) whose following vowel contexts were carefully controlled.

**Table 2.** Ten tokens in SP tests in relation to following vowel conditions

Vowel height and backness	Front vowels	Central vowels	Back vowels
<b>High Vowels</b>			
Familiar items	<i>read</i>		<i>roof</i>
Unfamiliar items	<i>ring</i>		<i>route</i>
<b>Mid Vowels</b>			
Familiar items	<i>rain</i>		<i>road</i>
Unfamiliar items	<i>red</i>	<i>rush</i>	<i>rope</i>
<b>Low Vowels</b>			
Familiar items		<i>rice</i>	
Unfamiliar items			

*Note.* *Road* was tested twice in both the CP and SP tests.



**Table 3.** Fifteen tokens in CP tests in relation to following vowel conditions

Vowel height and backness	Front vowels	Central vowels	Back vowels
High Vowels			
Familiar items	<i>rink, reef</i>		<i>rule, room</i>
Unfamiliar items	<i>reach</i>		<i>rude</i>
Mid Vowels			
Familiar items	<i>race, rent</i>		<i>road, wrong</i>
Unfamiliar items	<i>rate</i>	<i>rough</i>	<i>roll</i>
Low Vowels			
Familiar items		<i>Ryan, right</i>	
Unfamiliar items			

*Note.* *Road* was tested twice in both the CP and SP tests.

In total, 1,470 tokens (i.e., 15 words  $\times$  49 students  $\times$  2 test sessions = 1,470 tokens) were produced at both the pre- and posttest sessions. These words are listed in Table 3.

## NS Baseline

For comparison purposes, 10 NSs (males,  $n = 6$ ; females,  $n = 4$ ) were asked to complete the CP and SP tests following the same procedure as the Japanese learners. At the time of the study, all of them were undergraduate students studying at an English-speaking university in Montreal and spoke North American English as their L1 (age:  $M = 21.54$ ,  $SD = 1.64$ ). Their performance was used to demonstrate to what degree NS tokens differ from those produced by Japanese learners. In total, they generated 150 tokens (i.e., 15 words  $\times$  10 NSs = 150 tokens) for the CP tests and 100 words (i.e., 10 words  $\times$  10 NSs = 100 tokens) for the SP tests.

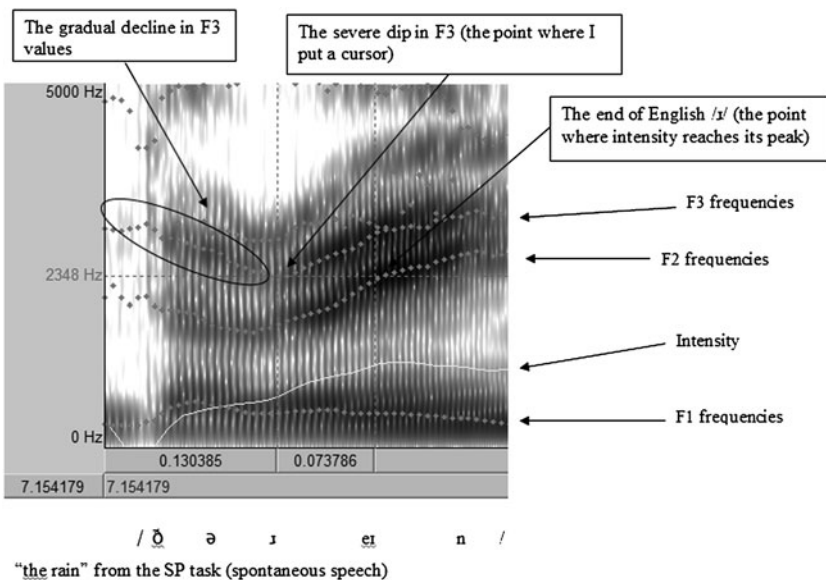
## Acoustic Analyses

Acoustic analyses were conducted on the primary acoustic property of /ɹ/—that is, F3 values—in all 2,700 tokens (i.e., 2,450 words from 49 learners + 250 words from 10 NSs) to assess in depth to what degree Japanese learners exhibited gains resulting from FFI with and without EI in comparison with the NS baseline. As in Saito and Lyster's (2012) study, this study followed Flege et al.'s (1995) procedure to acoustically analyze spectral aspects of natural /ɹ/ tokens (i.e., F3 values) elicited

from a variety of production tasks. All speech tokens were analyzed via the speech analysis software Praat (Boersma & Weenik, 2009).

The beginning of word-initial English /ɪ/ was first identified via spectrographic representations and wave forms of the speech tokens in conjunction with the onset of the energy for all three formants. For /ɪ/ embedded in continuous speech (i.e., the SP tokens), every effort was made to find the beginning of word-initial /ɪ/. Given that English /ɪ/ exhibits relatively low F3 values compared to other vowel and consonant sounds in the English phonetic system, as a reliable cue, the author carefully located the end of gradual decline in F3 values (for details of acoustic properties of /ɪ/, see Espy-Wilson, 1992; Espy-Wilson et al., 2000). The cursor was placed on the severe dip in F3, from which F3 values again start to increase toward following vowels, and the auto-correlation method of linear predictive coding analysis was used to measure F3 values of /ɪ/ at this point.<sup>8</sup> To reduce spectra variations due to talker variability (e.g., gender and length of vocal tract), normalization of the raw acoustic values was conducted by converting Hz into Bark (see Boersma & Weenik, 2009, for details of normalization procedures). The acoustic analysis procedure is visually summarized in Figure 2.

An intercoder reliability session was administered to check the construct validity of the acoustic analysis procedure previously described. Five speech stimuli (i.e., three CP tokens and two SP tokens) were randomly chosen from each of the participants. The subset of 245 tokens



**Figure 2.** Procedure for the acoustic analysis of English /ɪ/.

(i.e., 49 Japanese learners  $\times$  5 tokens = 245 tokens)—of the 2,700 total tokens—accounted for roughly 10% of the entire dataset. The author and one experienced phonetician conducted acoustic analyses separately. Significant positive correlations were found between coders for the entire data set,  $r(244) = .774, p < .001$ ; the CP tokens,  $r(146) = .738, p < .001$ ; and the SP tokens,  $r(97) = .834, p < .001$ . In conjunction with the high intercoder reliability in all contexts, the author analyzed the remaining dataset.

## Endpoint Interview

To investigate the role that students' perceptions of FFI might have played in determining its effectiveness, the 35 learners in the experimental groups were individually interviewed (FFI+EI group,  $n = 17$ ; FFI-only group,  $n = 18$ ) after they completed the posttest. After receiving an explanation about the primary purpose of the current project (i.e., teaching how to produce and perceive /ɪ/ in meaning-oriented classrooms), participants were asked to estimate in a retrospective manner what proportion of their focus had been on meaning (i.e., English argumentative skills) and on form (i.e., intelligible pronunciation of /ɪ/) during the FFI treatment.

## RESULTS

### Pre- and Posttest Data

The pre- and posttest scores were sorted out according to two lexical contexts: (a) familiar items and (b) unfamiliar items. Each participant's F3 values were then averaged to obtain a group mean for each of the two variables: test type and following vowel backness. Separate three-factor ANOVAs (Group  $\times$  Time  $\times$  Test; Group  $\times$  Time  $\times$  Backness) were performed to find any statistically significant differences between pre- and posttest sessions (i.e., within-group comparison) and among the three groups at the posttest sessions (i.e., between-group comparison). Finally, any significant change in F3 values was interpreted in accordance with the benchmark of NS perceptions of /ɪ/ established in Saito (2011)<sup>9</sup>:

1. *Good English* /ɪ/ (F3 = 12.97–13.42 Bark [2,028–2,174 Hz])
2. *Poor English* /ɪ/ (F3 = 13.72–14.42 Bark [2,283–2,529 Hz])
3. *Hybrid exemplars* (F3 = 14.50–15.14 Bark [2,550–2,797 Hz])
4. *English* /ɪ/ (F3 > 15.29 Bark [2,850 Hz])

The scores for the 10 NSs who took the same tests once are also reported for reasons of comparison. An alpha level was set at  $p < .05$  for

all statistical analyses. Cohen's  $d$  was also calculated to measure the magnitude of instructional effectiveness between two contrast groups of means. 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 ( $1.00 \leq d$ ).

**Pretest Data.** To check the preexisting difference in F3 values between the three groups—FFI-only, FFI+EI, and control—according to two types of lexical items (i.e., familiar vs. unfamiliar tokens), the pretest scores were submitted to three-factor ANOVAs for each of the test and following vowel conditions (Group  $\times$  Lexis  $\times$  Test; Group  $\times$  Lexis  $\times$  Backness). Neither main effects of group nor lexis were found significant in any contexts,  $p = .300$ – $.800$ . This indicates that any changes in the experimental groups were attributable to neither group nor lexical difference at the onset of the study.

*Familiar items: Task types.* A three-factor repeated-measures ANOVA, with test (i.e., the CP and SP tests) and time (i.e., pre- and posttests) as the repeated measure and group (i.e., FFI-only, FFI+EI, and control) as the between-group factor, found overall main effects for test,  $F(1,46) = 7.137$ ,  $p < .001$ . The result indicated that the participants in the current study had slightly more difficulty in the SP test ( $M = 14.47$  Bark,  $SD = 3.25$  Bark) than in the CP test ( $M = 14.34$  Bark,  $SD = 3.31$  Bark).

The ANOVA also found overall main interaction effects for group and time,  $F(2, 46) = 24.298$ ,  $p < .001$ . Tukey-Kramer post hoc pairwise comparisons were employed to find the source of the significance, revealing several interesting patterns. First, the members of the FFI+EI group exhibited significant improvement in their performance of English /ɪ/ from hybrid exemplars ( $M = 14.68$  Bark,  $SD = 0.68$  Bark) to good exemplars ( $M = 13.53$  Bark,  $SD = 0.92$  Bark),  $p < .001$ , with large effects,  $d = 1.42$ . They also outperformed the control group ( $M = 14.71$  Bark,  $SD = 0.52$  Bark) at the time of the posttest sessions,  $p < .001$ , with large effects,  $d = 1.58$ . Second, the FFI-only group showed significant change over time from hybrid exemplars ( $M = 14.71$  Bark,  $SD = 0.76$  Bark) to poor exemplars ( $M = 14.07$ ,  $SD = 0.95$ ),  $p < .001$ , and outperformed the control group ( $M = 14.71$  Bark,  $SD = 0.52$  Bark),  $p = .045$ , with medium effects ( $d = 0.74$  for the within-group difference and  $d = 0.81$  for the between-group difference).

*Familiar items: Vowel backness.* A three-way ANOVA was conducted: Group (i.e., FFI-only, FFI+EI, and control)  $\times$  Time (i.e., pre- and posttest)  $\times$  Backness (i.e., singletons with front, central, and back vowels). The ANOVA found significant overall main effects for backness,  $F(1, 46) = 97.233$ ,  $p < .001$ . The learners produced /ɪ/ with different F3 values according to its following vowel backness in the following order: (a) front vowels ( $M = 14.61$  Bark,  $SD = 0.86$  Bark)  $>$  (b) central vowels ( $M = 14.32$  Bark,  $SD = 1.00$  Bark)  $>$  (c) back vowels ( $M = 14.03$  Bark,  $SD = 0.89$  Bark).

The ANOVA revealed overall main interaction effects for group and time,  $F(2, 46) = 24.612$ ,  $p < .001$ . According to Tukey-Kramer post hoc

pairwise comparisons, the members of the FFI+EI group significantly changed their performance of /ɪ/ from hybrid exemplars ( $M = 14.67$  Bark,  $SD = 0.79$  Bark) to good exemplars ( $M = 13.43$  Bark,  $SD = 0.92$  Bark),  $p < .001$ , with large effects,  $d = 1.45$ . They outperformed the control group ( $M = 14.67$  Bark,  $SD = 0.58$  Bark),  $p < .001$ , with large effects,  $d = 1.67$ , and the FFI-only group ( $M = 14.04$  Bark,  $SD = 1.01$  Bark),  $p < .05$ , with medium effects,  $d = 0.63$ , at the time of the posttest sessions. The FFI-only group showed significant change over time from hybrid exemplars ( $M = 14.69$  Bark,  $SD = 0.83$  Bark) to poor exemplars ( $M = 14.04$  Bark,  $SD = 1.01$  Bark) with medium effects,  $d = 0.71$ , but did not significantly outperform the control group at the time of the posttest sessions,  $p > .05$ .

*Unfamiliar tokens: Task type.* A three-way ANOVA (Group  $\times$  Time  $\times$  Task) found significant overall interaction effects for group and time,  $F(2, 46) = 28.025$ ,  $p < .001$ . According to Tukey-Kramer post hoc pairwise comparisons, the members of the FFI+EI group exhibited significant improvement in their performance of /ɪ/ over time from hybrid exemplars ( $M = 14.68$  Bark,  $SD = 0.76$  Hz) to good exemplars ( $M = 13.55$  Bark,  $SD = 0.92$  Bark),  $p < .001$ , with large effects,  $d = 1.32$ . They also outperformed the control group ( $M = 14.66$  Bark,  $SD = 0.64$  Bark),  $p < .001$ , with large effects,  $d = 1.40$ , at the time of the posttest sessions. The FFI-only group demonstrated significant change over time from hybrid exemplars ( $M = 14.63$  Bark,  $SD = 0.71$  Bark) to poor exemplars ( $M = 14.27$  Bark,  $SD = 0.87$  Bark),  $p < .001$ , with small effects,  $d = 0.43$ , but did not outperform the control group at the time of the posttest sessions,  $p > .05$ .

*Unfamiliar tokens: Vowel backness.* A three-way ANOVA (Group  $\times$  Time  $\times$  Backness) found significant overall interaction effects for group and time,  $F(2, 46) = 28.033$ ,  $p < .001$ . According to Tukey-Kramer post hoc pairwise comparisons, the members of the FFI+EI group revealed significant improvement in their performance of /ɪ/ over time from hybrid exemplars ( $M = 14.68$  Bark,  $SD = 0.80$  Bark) to good /ɪ/ exemplars ( $M = 13.52$  Bark,  $SD = 0.98$  Bark),  $p < .001$ , with large effects,  $d = 1.29$ , and they outperformed the control group ( $M = 14.68$  Bark,  $SD = 0.68$  Bark),  $p < .001$ , with large effects,  $d = 1.37$ , at the time of the posttest sessions. The FFI-only group demonstrated significant change over time from hybrid exemplars ( $M = 14.63$  Bark,  $SD = 0.76$  Bark) to poor exemplars ( $M = 14.10$  Bark,  $SD = 0.96$  Bark),  $p < .001$ , with small effects,  $d = 0.58$ , but did not outperform the control group at the time of the posttest sessions,  $p > .05$ . A summary of all relevant results appears in Table 4.

## NS Baseline

The descriptive statistics showed that 10 NSs generally produced /ɪ/ with relatively low F3 values for both familiar items ( $M = 11.51$  Bark,

**Table 4.** Summary of the significant changes in F3 values

Group	Lexical contexts	Variables	Within-group comparisons (pre- → posttests)	Between-group comparisons (vs. control/FFI-only at posttests)
FFI+EI group ( <i>n</i> = 17)	Familiar items	Test types	Large effects ( <i>M</i> = 14.68 → 13.53 Bark, <i>d</i> = 1.42)	Large effects (vs. control, <i>d</i> = 1.58)
		Vowel backness	Large effects ( <i>M</i> = 14.67 → 13.43 Bark, <i>d</i> = 1.45)	Large effects (vs. control, <i>d</i> = 1.61) Medium effects (vs. FFI-only, <i>d</i> = 0.63)
	Unfamiliar items	Test types	Large effects ( <i>M</i> = 14.68 → 13.55 Bark, <i>d</i> = 1.32)	Large effects (vs. control, <i>d</i> = 1.40)
		Vowel backness	Large effects ( <i>M</i> = 14.68 → 13.52 Bark, <i>d</i> = 1.29)	Large effects (vs. control, <i>d</i> = 1.37)
FFI-only group ( <i>n</i> = 18)	Familiar items	Test types	Medium effects ( <i>M</i> = 14.71 → 14.07 Bark, <i>d</i> = 0.74)	Medium effects (vs. control, <i>d</i> = 0.81)
		Vowel backness	Medium effects ( <i>M</i> = 14.69 → 14.04 Bark, <i>d</i> = 0.71)	<i>n.s.</i>
	Unfamiliar items	Test types	Small effects ( <i>M</i> = 14.63 → 14.27 Bark, <i>d</i> = 0.43)	<i>n.s.</i>
		Vowel backness	Small effects ( <i>M</i> = 14.63 → 14.10 Bark, <i>d</i> = 0.58)	<i>n.s.</i>

*SD* = 1.45 Bark) and unfamiliar items (*M* = 11.58 Bark, *SD* = 1.58 Bark). A matched-paired *t* test found no significant differences between the two lexical contexts, *p* = .326. As for the two affecting variables, a set of one-factor repeated ANOVAs found that, although no difference was found in their performance between the CP tests and SP tests, NSs tended

to produce lower F3 values for /ɪ/ preceding central and back vowels ( $M = 11.72$  Bark,  $SD = 1.70$  Bark)<sup>10</sup> than front vowels ( $M = 11.45$  Bark,  $SD = 1.58$  Bark),  $F(2, 18) = 4.923$ ,  $p = .019$ .

## Post Hoc Analyses

According to the final interview, the learners in both groups were split quite evenly between those who claimed to have focused on form and those who focused on meaning. The analysis of individual self-report scores identified (a) 18 learners with relative focus on meaning ( $n = 10$  for the FFI+EI group, and  $n = 8$  for the FFI-only group), (b) 6 learners with equal focus on meaning and form ( $n = 2$  for the FFI+EI group, and  $n = 4$  for the FFI-only group), and (c) 11 learners with relative focus on form ( $n = 5$  for the FFI+EI group, and  $n = 6$  for the FFI-only group).

## DISCUSSION

Given that Japanese learners generally produced hybrid exemplars (F3 values = 14.50–15.14 Bark) at the onset of the study, compared to the NS baseline (F3 values = 11.45–11.79 Bark), in this section, the degree to which FFI alone was facilitative of learners' L2 pronunciation development of /ɪ/—in both familiar and unfamiliar lexical contexts—is reexamined as assessed by measures that had been improved since the Saito and Lyster (2012) study (e.g., by including familiar and unfamiliar items in pre- and posttest materials). Subsequently, the discussion examines to what degree adding EI to FFI lessons enhances the acquisitional value of the overall instructional treatment.

### Reexamining FFI Effectiveness

For the first research question, which concerns the generalizability and magnitude of FFI-only effectiveness, the results of this study pointed to slightly different patterns of gain according to lexical contexts. For familiar items, the FFI-only group members changed their F3 values from 14.60–14.70 Bark to  $F3 = 14.00$ –14.20 Bark with medium effects between pre- and posttest sessions. Additionally, their posttest scores were greater than those of the control group with medium effects under the task type condition. For unfamiliar items, the amount of F3 change was small, and they did not outperform the control group in any contexts; that is, the FFI-only group generally enhanced their performance from



hybrid exemplars to poor English /ɹ/ exemplars particularly in the familiar lexical items (i.e., medium effects, some between-group difference) versus the unfamiliar items (i.e., small effects, no between-group difference). The results of the current study, with beginner-intermediate Japanese EFL learners in Japan, echoed those of the original study, with intermediate Japanese ESL learners in Canada.

Unlike in traditional decontextualized instruction methods (e.g., audio-lingual methods), in the current study—as well as in Saito and Lyster (2012)—FFI was integrated into meaning-oriented lessons, and this encouraged learners to focus on both form and meaning simultaneously (i.e., out of 18 learners, 12 learners self-reported either more focus on meaning than form or equal focus on both form and meaning during FFI lessons). In line with similar FFI studies in L2 morphosyntax development, this type of FFI appeared to impact learners' developing L2 phonological system equally at both a controlled- and spontaneous-speech level, irrespective of the following vowel contexts. It is important to note, however, that the magnitude of FFI effectiveness was small to medium, and its generalizability to unfamiliar lexical contexts was limited to some degree. This indicates that FFI delivered at a lexical level may not be sufficient to trigger L2 learners' phonemic and phonetic noticing and awareness, especially in the case of the relatively difficult segmental acquisition of /ɹ/ by Japanese learners. That is, these learners may need either more exposure to exemplars or other kinds of modified input beyond FFI to establish a new phonetic category for /ɹ/ in their long-term memory representation.

### **Adding EI to FFI Lessons**

With respect to the second research question, which asked whether and to what degree providing EI enhances the generalizability and magnitude of FFI effectiveness, the current study found that the FFI+EI group members generally changed their F3 values over time from 14.60–14.70 Bark to 13.40–13.50 Bark with large effects for both familiar and unfamiliar items ( $d = 1.30$ – $1.40$ ). In other words, the FFI+EI group not only demonstrated considerable improvement from hybrid exemplars to good exemplars but also transferred the instructional gain from familiar lexical contexts to unfamiliar lexical contexts. It is important to note that, out of four possible contexts (i.e., 2 lexical factors [familiar and unfamiliar items]  $\times$  2 affecting variables [test types and vowel backness]), the FFI+EI group outperformed the control group with large effects ( $d = 1.50$ – $1.65$ ) in all contexts and the FFI-only group with small effects under one condition (i.e., familiar items in the vowel backness category). According to the interview data, EI did not interfere with the learners'

simultaneous focus on form and meaning (out of 17 learners, 12 self-reported either more focus on meaning than form or equal focus on both form and meaning during FFI lessons).

As predicted earlier, the advantage of adding EI to FFI treatment might mainly be due to several factors. First, the instructor's exaggerated model pronunciation of only /ɪ/ (and /l/) would directly help the learners notice the perceptual difference between the new sound (/ɪ/) and its L1 counterpart (i.e., the Japanese tap /ɾ/). Additionally, the learners were explicitly taught the primary articulatory configurations for /ɪ/ (i.e., constrictions in lips, palatal regions, and pharynx), which might have also promoted their increased awareness of the new sound form in a complementary fashion (Catford & Pisoni, 1970).

Explicit phonetic information was implemented before FFI lessons so that the beginner-intermediate learners in the current study could fully use their limited attentional resources to attend to the phonetic unit of L2 input under no communicative pressure. In this way, these learners with high sensitivities to sound-sized units of L2 phonological information could make the best of the following FFI activities, in which the new sound form was embedded at a lexical level. Namely, they practiced the target sound feature via a number of communicative activities to (a) proceduralize more targetlike representations (i.e., the large gain was found across different tasks and following vowel conditions) and (b) generalize the newly acquired phonetic knowledge to unfamiliar lexical contexts (the large effects were apparent not only in familiar items but also in unfamiliar items). The results would support the view that EI may be necessary with respect to linguistic features that learners would otherwise have tremendous difficulties in noticing through mere exposure to L2 input (Ellis, 2002; Spada & Lightbown, 2008; VanPatten, 2004).

## CONCLUSION AND FUTURE DIRECTIONS

The current study was designed to reexamine generalizability and magnitude of FFI effectiveness as well as to test the pedagogical capability of EI to promote learners' phonetic-level restructuring in the context of L2 pronunciation development of /ɪ/ by Japanese learners. The results revealed that the learners who received FFI with and without EI demonstrated two different types of L2 phonological development. On the one hand, the learners in the FFI-only group demonstrated medium improvement, especially in familiar lexical contexts. Their gains, however, were not fully transferred to unfamiliar lexical contexts. This finding implies that, without remedial techniques such as EI, L2 learners tend to decode L2 phonological information at a lexical level as their default interlanguage strategy, and their access to the relevant mental representations

is subject to the influence of lexical items occurring in the instructional materials and is, thus, context specific.

On the other hand, the learners in the FFI+EI group, who started with the noticing of perceptual aspects of /ɪ/ via EI before FFI, showed (a) large gains (i.e., change from hybrid exemplars to good /ɪ/ exemplars) across different tasks and following vowel conditions and (b) the ability to generalize their gains to new lexical contexts beyond the instructional materials. This result indicates that these learners established strong context-invariant representations in the underlying system with various levels of processing abilities (CP vs. SP, various following vowel conditions, etc.). In short, the results of the current study showed the relative impact of phonetically driven L2 phonological learning (FFI+EI) over lexically driven L2 phonological learning (FFI-only), especially in the case of the difficult segmental acquisition of /ɪ/ by Japanese learners.

Finally, the findings of the current study provide several pedagogical recommendations for teaching L2 segmental sounds in relation to the FFI model developed by Lyster and Ranta (Lyster, 2007; Ranta & Lyster, 2007): noticing → awareness → practice. At the noticing phase, some isolated intervention—such as EI—may be necessary to push L2 learners to (a) attend to sound-sized units of L2 phonological information and (b) notice the perceptual difference between a new sound and its L1 counterpart. This phonetic-level noticing could be a first step toward the restructuring of existing representations and the establishment of a new phonetic category in a learner's representational system (Flege, 1995, 2003, 2009). At the awareness stage, L2 learners should then be given communicative tasks in which they can further process the target sound either receptively or productively in meaningful lexical contexts, to help learners develop and internalize the phonetic representation. At the practice phase, L2 learners should be encouraged to repetitively practice the target sound feature in communicatively authentic contexts to proceduralize their targetlike phonetic representations (i.e., a transition from effortful to automatic use of phonetic knowledge). Given that the corrective force of pronunciation-focused recasts is quite salient to L2 learners (i.e., L2 learners' repair rate following pronunciation-focused recasts is reported at around 80% in a wide range of classrooms; see Lyster, 1998; Sheen, 2006), teachers could use this approach to push their students to repair their nontargetlike production of target sounds at a lexical level.

To conclude, I would like to emphasize a strong call for more L2 speech acquisition research of this kind to further examine at what point EI, focused tasks, and CF should be introduced to L2 learners according to the differential level of their developing representation and processing abilities. To this end, one suggestion for future studies could be to adopt not only production measures but also perception measures because change in the perception phase entails change in a learner's represen-

tational system and thus precedes change in the production phase (Bradlow et al., 1997; Flege, 1995, 2003, 2009). Given that both the current study and the Saito and Lyster (2012) study focused only on change in the production phase, despite some efforts made at each phase (i.e., adopting not only CP but also SP), it is still possible that the learners could have consciously and carefully produced /ɪ/ drawing on their explicit articulatory knowledge (i.e., monitoring) even without establishing or developing a new representation. In other words, it would be intriguing to implement a wide range of perception and production tests simultaneously, to examine (a) which combination of FFI techniques actually impacts the developing L2 system at the initial stages of L2 development, which could be mainly measured via perception tests; and (b) to what degree they proceduralize the newly acquired phonetic knowledge at later stages of L2 development, which could be mainly measured via production tests.

*Received 9 June 2011*

*Accepted 29 September 2011*

*Final version received 17 October 2011*

## NOTES

1. To measure the effectiveness and durability of FFI, the posttest sessions in the current study took place two weeks after instruction. This testing interval could be considered as short delayed, rather than immediate, according to the FFI research standards in L2 grammar studies (e.g., Spada & Tomita, 2010). It has been shown that the integration of novel phoneme sequences into the mental lexicon takes around a week rather than occurring immediately after the exposure (Gaskell & Dumay, 2003).

2. Given that the FFI+EI teacher also taught one FFI-only class and that the control group teacher also taught FFI, a reviewer pointed out that it might have been challenging for the teachers to turn off the FFI or EI aspects of the teaching they had been trained to provide. To avoid this possible problem, several efforts were made in the current project. Not only did the researcher observe and videotape the instructional treatments but the teachers were also always reminded of the types of FFI treatment before the beginning of each class. Additionally, it was crucial to provide the participating teachers with a sufficient amount of training (i.e., 4 hr) and to ensure their understanding of the purpose of the entire research.

3. The same male instructor administered EI for all EI classes due to a schedule conflict. To check the nature and comparability of the instructional treatment, his production of /ɪ/ was elicited in the citation form of five words (i.e., *reef*, *race*, *rice*, *roof*, and *wrong*) and then was acoustically analyzed according to the procedure described in the Acoustic Analyses section. The primary acoustic parameter of his /ɪ/ (i.e., F3) ranged from 1,644 Hz (in *wrong*) to 1,926 Hz (in *reef*), which concurs with a typical F3 coverage of North American English /ɪ/ across sex and phonetic positions (i.e., 1,300 Hz–1,950 Hz; see Espy-Wilson et al., 2000).

4. Catford and Pisoni (1970) showed that teaching relevant articulatory gestures helped learners perceive new sounds and concluded that “‘exotic sounds’ can generally be more readily and unerringly identified after one has learned to produce them” (p. 481).

5. The combination of these articulatory configurations is categorized as “tip-up bunched /ɪ/,” which includes three crucial constrictions for producing /ɪ/ (i.e., lips, palatal

regions, and pharynx), and thus covers a range of variants between “tip-up retroflex /ɹ/” and “tip-down bunched /ɹ/” (Espy-Wilson et al., 2000, p. 345).

6. For pedagogic standpoints, Bradlow (2008) commented that “the lip rounding feature of English /ɹ/ production can be a useful characteristic to stress when teaching English pronunciation” (p. 292).

7. This time-pressure approach is not without problems. For example, Ellis (2002) himself acknowledged that “free-production tasks make it difficult but not impossible for learners to perform on the basis of explicit knowledge” (p. 234). In a similar manner, DeKeyser (2003) pointed out that the use of time pressure “merely made the use of explicit knowledge more difficult, and not impossible” (p. 326).

8. The beginning of following vowels (i.e., the endpoint of /ɹ/) was also roughly estimated by checking where the intensity reached its peak (see Figure 2).

9. The 95% confidence intervals of the F3 range were calculated on the basis of the impressionistic judgments of 20 NS listeners, who rated a continuum of natural /ɹ/-/l/ tokens by Japanese learners on a nine-point rating scale (point ranges listed in parentheses): (a) *Good English* /ɹ/ ( $1 \leq x \leq 2.5$ ), (b) *poor English* /ɹ/ ( $2.5 < x \leq 4$ ), (c) *hybrid exemplars*—that is, neither English /ɹ/ nor English /l/—( $4 < x < 6$ ), (d) *poor English* /l/ ( $6 \leq x < 7.5$ ), and (e) *good English* /l/ ( $7.5 \leq x \leq 9$ ). For similar results, see Flege and colleagues (1995).

10. The Tukey test did not find any significant difference between singletons with central vowels and back vowels.

## REFERENCES

- 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: Benjamins.
- Boersma, P., & Weenik, D. (2009). Praat: Doing phonetic by computer [Computer software]. Retrieved December 2009, from <http://www.praat.org>.
- Bradlow, A. (2008). Training non-native language sound patterns. In J. Hansen & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 287–308). Amsterdam: 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–n2085.
- Bradlow, A., Pisoni, D., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/. *Journal of the Acoustical Society of America*, *101*, 2299–2310.
- Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (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. L., Best, C. T., & Tyler, M. D. (2011b). Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. *Applied Psycholinguistics*, *32*, 51–67.
- Campbell, F., Gick, B., Wilson, I., & Vatikiotis-Bateson, E. (2010). Spatial and temporal properties of gestures in North American English /ɹ/. *Language and Speech*, *53*, 49–69.
- Catford, J., & Pisoni, D. (1970). Auditory vs. articulatory training in exotic sounds. *Modern Language Journal*, *54*, 447–481.
- Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for learners of English to speakers of other languages*. New York: Cambridge University Press.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody and the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–201.
- DeKeyser, R. (2003). Implicit and explicit learning. In M. Long & C. Doughty (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford: Blackwell.
- DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: 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.
- Diehl, R., Lotto, A., & Holt, L. (2004). Speech perception. *Annual Review of Psychology*, 55, 149–179.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In M. Long & C. Doughty (Eds.), *Handbook of second language acquisition* (pp. 257–310). Oxford: Blackwell.
- Elliott, A. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 80, 95–108.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? *Studies in Second Language Acquisition*, 24, 223–236.
- Espy-Wilson, C. (1992). Acoustic measures for linguistic features distinguishing the semi-vowels /wjr/ in American English. *Journal of Acoustical Society of America*, 92, 736–751.
- Espy-Wilson, C., Boyce, S., Jackson, M., Narayanan, S., & Alwan, A. (2000). Acoustic modeling of American English /r/. *Journal of the Acoustical Society of America*, 108, 343–356.
- Fernandez, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition*, 30, 277–305.
- Flege, J. (1989). Chinese subjects' perception of the word-final English /t-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, 86, 1684–1697.
- 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. (2009). Give input a chance! In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 175–190). Bristol, UK: Multilingual Matters.
- 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., Takagi, N., & Mann, V. (1995). Japanese adults learn to produce English /ɹ/ and /l/ accurately. *Language and Speech*, 38, 25–55.
- Gaskell, G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, 89, 105–132.
- 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 /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America*, 125, 469–479.
- Henry, N., Culman, H., & VanPatten, B. (2009). More on the effects of explicit information in instructed SLA. *Studies in Second Language Acquisition*, 31, 559–575.
- 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.
- Iverson, P., Kuhl, P. K., 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.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science*, 97, 11850–11857.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831–843.
- Larson-Hall, J. (2006). What does more time buy you? Another look at the effects of long-term residence on production accuracy of English /r/ and /l/ by Japanese speakers. *Language and Speech*, 49, 521–548.



- Lively, S., Logan, J., & Pisoni, D. (1993). Training Japanese listeners to identify English /ɪ/ and /I/. II: The role of phonetic environments and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242–1255.
- Lotto, A. J., Sato, M., & Diehl, R. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/. In J. Slifka, S. Manuel, & M. Matthies (Eds.), *From sound to sense: 50+ years of discoveries in speech communication* (pp. C381–C386). Cambridge, MA: Research Laboratory of Electronics at MIT.
- Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48, 183–218.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: Benjamins.
- 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.
- 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). New York: Cambridge University Press.
- Saito, K. (2011). *Effects of FFI on L2 phonological development of /ɪ/ by Japanese learners of English* (Unpublished doctoral dissertation). McGill University, Montreal, Canada.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɪ/ by Japanese learners of English. *Language Learning*, 62, 595–633.
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10, 361–392.
- Sheldon, A., & Strange, W. (1982). The acquisition of /ɪ/ and /I/ 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., & Lightbown, P. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42, 181–207.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, 263–308.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–484). Mahwah, NJ: Erlbaum.
- Trofimovich, P., & Gathbonton, E. (2006). Repetition and focus on form in L2 Spanish word processing: Implications for pronunciation instruction. *Modern Language Journal*, 90, 519–535.
- VanPatten, B. (2004). *Input processing and grammar instruction in second language acquisition*. Westport, CT: Ablex.
- 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: Benjamins.
- Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E., Kawakatsu, M., Tohkura, Y., & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *Neuroimage*, 46, 226–240.