The Journal of Asia TEFL

Shadowing for Pronunciation Development: Haptic-Shadowing and IPA-Shadowing

Yo Hamada

Akita University

Shadowing was introduced to foreign language teaching in the 1990s and since then it has been researched in EFL teaching contexts. Shadowing has been used as a listening task, and its effectiveness on listening comprehension has been acknowledged. However, its effect on pronunciation development remains unclear. With this unresolved issue in mind, I have attempted to develop shadowing-based pronunciation teaching methods by combining shadowing with activities for pronunciation skill development, namely haptic-shadowing and IPA-shadowing. A total of 58 Japanese second-year university students participated in the experiment (29 for haptic-shadowing and 29 for IPA-shadowing). A total of 15 lessons were given to each group. The results show that the haptic-shadowing group statistically improved on all the three features of comprehensibility, segmental features, and suprasegmental features, while the IPA-shadowing group improved on comprehensibility and segmental features. Additionally, an exploratory factor analysis was conducted to examine learners' perceptions toward the two types of shadowing. Learners have a more positive and fresh image of haptic-shadowing and IPA-shadowing.

Keywords: Shadowing, pronunciation, Haptic, IPA, TPR

Introduction

Ever since *shadowing* was introduced to foreign language teaching more than two decades ago (Tamai, 1992), it has been researched in EFL teaching contexts. Researchers would now agree that shadowing is effective for improving learners' listening comprehension skills, but the positive efficacy of shadowing for speaking skill development remains questionable (Hamada, 2016a). However, because learners do produce speech sounds – by replicating what they hear when shadowing – shadowing is assumed to benefit overall speaking skill development (Hamada, 2016b). Though limited, research hints that shadowing may have a positive impact on pronunciation development (Mori, 2011). Pronunciation has caught researchers' attention and the effectiveness of pronunciation teaching has been recognized (Derwing, 2012; Lee, Jang, & Plonsky, 2014; Saito, 2012), although insufficient attention has been paid to pronunciation in classroom-oriented research (Baker, 2014). Given these circumstances, this study explores the possibility of shadowing as a technique for pronunciation development. The underlying purpose of this two-study article is to develop shadowing-based teaching techniques to improve learners' pronunciation in classroom. In the first study, the effectiveness of the techniques on pronunciation development will be examined, and in the second study, the learners' perceptions of the techniques will be explored.

Literature Review

Shadowing and Pronunciation

A brief overview of shadowing will be provided below (for more details, see Hamada, 2016a, 2016c). L1 Shadowing has been researched since around 1970s (e.g., Carey, 1971), and it was also used in the early stages of training interpreters (Lambert, 1988). Later it was imported to EFL contexts, an initiative spearheaded by Tamai (1992). Shadowing is defined as "a paced, auditory tracking task which involves the immediate vocalization of auditorily presented stimuli" (Lambert, 1992, p. 266), and research on the topic has spread in Japan. Case studies have been conducted mostly in secondary schools and universities (e.g., Hamada, 2016c; Tamai, 1992, 1997). The empirical research has assured that shadowing improves bottom-up listening skills and listening comprehension skills.

Shadowing looks similar to other repetition based techniques, but the cognitive process required by shadowing is different from that of other activities. Since there is little time lag between when learners hear the audio input and when they replicate it, shadowing is called an on-line task (Shiki, Mori, Kadota, & Yoshida, 2010). In contrast, repetition is an off-line task because learners wait until one chunk is read out in full before repeating it. Being an on-line activity, learners' listening and their vocalization of what they listen to occur simultaneously in shadowing, while the vocalization comes after listening to each chunk (Hamada, 2016a; Kadota, 2007).

While a certain amount of research on shadowing and listening has been published, only limited numbers of studies were published in terms of shadowing and speaking (e.g., Bovee & Stewart, 2009; Hsieh, Dong, & Wang, 2013; Iino, 2014; Mori, 2011). These case studies have implied that shadowing may be beneficial for learners' pronunciation development, but the research has not arrived at a general conclusion.

In terms of pronunciation teaching, *intelligibility*, *comprehensibility*, and *accentedness* often catch researchers' attention. Intelligibility is "the extent to which a listener understands the productions of an L2 speaker" (Derwing, 2012, pp. 1-2), comprehensibility is how easy or difficult L2 speech is to understand, and accentedness is a rating of the degree of difference between an L2 speaker's productions and the local variety of the language in question. In real-world contexts, comprehensibility is "central to interlocutors' communicative success" (Isaacs & Trofimovich, 2012, p. 475). Also, learners may "reach a threshold of comprehensibility while still being fairly accented" (Saito, Trofimovich, & Isaacs, 2015, p. 458). Therefore, I would like to focus on comprehensibility in this exploratory study.

Additionally, there have been debates over which segmental features (consonants and vowels) and suprasegmental features (rhythm, stress, and intonation) are influenced by shadowing (Avery & Ehrilich, 1992). However, given that Japanese is quite different from English in both respects, this exploratory study should examine both.

In summary, shadowing has been used as a listening task, and its effectiveness on listening comprehension skills has been acknowledged. However, its effect on pronunciation development remains unclear. With this unresolved issue in mind, I have attempted to develop shadowing-based pronunciation methods by combining shadowing with activities for pronunciation skill development, focusing on comprehensibility and the segmental and suprasegmental features of English.

Development of New Shadowing Techniques

Conditions. Certain important realities must be considered when choosing activities that combine shadowing and pronunciation for effective use in a general language classroom. The first issue is learners' proficiency levels. Japanese university EFL learners are generally considered to fall somewhere between near-intermediate and intermediate speakers. As an example, the average score of the Japanese test takers of the TOEFL is relatively lower than those in other countries (ETS, 2015). On the CEFR, most Japanese English learners belong to the A level (Negishi, Tono, & Fujita, Y, 2012), which is the basic level. The

pronunciation techniques, then, should not require a high English proficiency; they must be within the learners' reach. The second issue is the situations of the teachers. It is realistically challenging to find enough time for pronunciation training in class. In junior and senior high schools, teachers are required to teach a tight and fixed curriculum authorized by the Ministry of Education, and in university, the most common English classes are English general education courses. A straightforward technique such as shadowing is readily adaptable in these contexts.

Additionally, there are at least two conditions that must be met in order to connect shadowing and another activity. First, because shadowing is already cognitively demanding, the other activity should require less mental effort to avoid cognitive overload. Second, because shadowing is an on-line activity, the other activity should also be an on-line activity in terms of cognitive processes in order to avoid theoretical confusion and learners' additional intellectual strain.

Ideas for shadowing and pronunciation. In this study, the author developed new shadowing methods: Haptic-Shadowing (H-SH), and International Phonetic Alphabet-Shadowing (IPA-SH). The background and details of these methods are explained as follows.

The basic meaning of haptic refers to *movement plus touch* (Teaman & Acton, 2013), and haptic in the language teaching context refers to "systematic hand movement across the visual field accompanying speech that typically terminates in a touch of some kind, like one hand touching the other" (Acton, Baker, Burri, & Teaman, 2013, p. 234). Therefore, haptic teaching is partly rooted in the theory of Total Physical Response (TPR) (Teaman & Acton, 2013), which holds that motor activity is a right-brain function preceding left-brain processing, and that it focuses on stress-free, low-anxiety-inducing activities (Brown, 2001).

Among the various haptic-teaching activities, Burri, Baker, and Acton (2016) propose an activity called the Rhythm Fight Club (RFC) technique, which is accompanied by teaching vocabulary. The basic steps are syllable awareness, word-stress awareness, and the physical experience of the stress (see https://vimeo.com/61195605). With RFC, learners "punch" the stressed syllable, which is good because teaching stress is highly recommended for teaching pronunciation (Hahn, 2004).

In an attempt to bring the two concepts of shadowing and the haptic movements together, the haptic technique RFC becomes a pronunciation training technique to teach stress. Literally, it would mean combining shadowing with movement and touch, thus haptic-shadowing. In H-SH, learners shadow what they are listening to and simultaneously they will 'punch' the most stressed syllable in the most stressed word. For unstressed syllables, they will make the punches shorter and weaker. For example, if they were working on the sentence "*Australia* has *Christmas* in *summer*", learners would shadow the sentence and air-punch the most stressed word, *summer*, and weakly air-punch *Australia* and *Christmas*. As the main focus of this activity is the acquisition of stress, which is a component of suprasegmental features, the learners' suprasegmental features should improve to some degree, in theory.

Example of H-SH

We	went to	Kyoto	on a school	trip	in June.
Weak	Weak	Strong	Weak	Weak	Weak

Next, because phonetic alphabets themselves are isolated and abstract, the IPA needs to be internalized and taught realistically (Grant, 2014). The advantages of teaching phonemic alphabets are summarized in Sypyra-Kzlowaska (2015), who noted that learners can notice the differences between English spelling and pronunciation, minimizing the potentially negative impact of the spelling to pronunciation; that it helps learners to internalize the sound system of English; and that they realize the gaps in their L1 and English knowledge. IPA provides visual reinforcement of auditory input.

I define IPA-SH as shadowing accompanied by scripts written in IPA. IPA-SH is thought to help learners' cognitive process while listening and speaking, raising their conscious awareness of segmental features. When learners work on normal shadowing, they perceive incoming sounds and then try to replicate those sounds. In the process, they often replace the actual incoming sounds with the sounds they know. For example, since the Japanese language does not have the sound /r/, novice Japanese learners cannot distinguish it well, so they replace /r/ with /l/ immediately after perceiving /r/ (unsuccessful phonological coding), and pronounce it as /l/. IPA-SH helps to fix this process, and it does so by visualizing the phonetic alphabet, which should help learners connect what they hear more directly with what they should pronounce.

Learner Perception

For teaching techniques to be effective in classrooms, they need not only to be supported by theory and empirical data, but also to take learners' feelings and psychology into consideration. Even the greatest method would not be effective if learners were reluctant to engage in it. In an attempt to approach shadowing differently, Hamada (2011) examined how learners perceive shadowing using a Semantic Differential (SD) method. The SD method is a technique originally designed to measure people's attitudes and feelings toward words, in which participants rate concept using a scale between two opposing adjectives. For example, for the word *democracy,* they have such choices as *good - - - - bad* and *weak - - - - strong*, and are supposed to choose between four rating scales in this case (Richards & Schmidts, 2010). In the questionnaire survey, Hamada conducted an exploratory factor analysis and found that learners generally have mixed perceptions of usefulness, plainness, and uselessness regarding shadowing. In other words, learners feel shadowing is generally useful, while admitting its simplicity. Building on this work, the present study also explores learners' perceptions of H-SH and IPA-SH.

Summary

The discussion above outlined the following three points: (1) shadowing has been researched primarily for listening skill development; (2) haptic-pronunciation teaching and the use of IPA, from a theoretical and situational standpoint, are useful pairings for combining shadowing with pronunciation teaching; and (3) learners' feelings about the methods must also be examined in order to see whether the teaching techniques will be appropriate for the classroom. Overall, this study will form the first step of a proposal for an effective shadowing-based pronunciation teaching technique, considering H-SH and IPA-SH in terms of both their effectiveness and the learners' psychological perspectives on the methods.

Study 1

Research Questions

In order to examine the general effects of H-SH and IPA-SH on pronunciation development, the research question is set as follows: 1) Will learners improve their comprehensibility, segmental productions, and suprasegmental productions by practicing H-SH? 2) Will learners improve their comprehensibility, segmental productions, and suprasegmental productions by practicing IPA-SH?

Methods

Participants

A total of 58 Japanese second-year university students (18-21 years old) participated in the experiment. The H-SH group had 29 students majoring in Engineering, and the IPA-SH group had 29 students majoring in Health Science. Both groups have students with mixed levels of English proficiency, mainly from near-intermediate to intermediate, based on the instructors' observation. They were highly motivated to study English in the lessons because they chose the class among several other choices, so few of them skipped the classes. The author received the participants' consent to use the data for research.

To assess the participants' performance, five raters were selected (see Table 1). This study is unique in that it selected only non-native speakers as raters, and this was done for various reasons. Now that English has become a global language, and the number of non-native speakers of English has surpassed that of native speakers (Crystal, 2003; Graddol, 2003), it is time to start accepting non-native speakers as qualified raters. In fact, studies have reported that native speakers and non-native speakers of English rate in a similar way, with non-native speakers sometimes being more lenient toward different varieties of English. Brooks (2013) compared native and non-native raters of the U.S. federal government speaking tests, finding no significant difference between the two; therefore, highly proficient non-native speakers should be considered qualified to act as raters. Saito and Shintani (2016) compared Canadian and multilingual Singaporean raters of 50 speeches made by Japanese English learners, and they found the multilingual Singaporean raters were more lenient about the participants' comprehensibility. Importantly, learners should aim to acquire comprehensible English that is not limited to English native speakers. This study does not focus on accentedness, but rather on comprehensibility, so there is no need to stick to the native speaker criterion. Given these, involving a mix of non-native speakers with various backgrounds was considered most apt for the current situation, so highly proficient non-native speakers were chosen for this study.

Materials

The experiments were conducted in the first 15 lessons of an EFL course, in which the concept of global varieties of English was taught. The course used an EFL textbook, *Understanding English Across Cultures* (Honna, Takeshita, & D'Angelo, 2012), which introduces and discusses English variations. In each lesson, the participants studied passages of approximately 450-550 words, and a 40-60 word passage was isolated that included important pronunciation features for Japanese EFL learners (i.e., /æ, f, v, θ, ð, w, l, ɹ/, as proposed in Saito, 2011). This passage was used for H-SH or IPA-SH.

In order to teach rules of stress, to the H-SH group, the course studied a section from another textbook (Avery & Ehrichs, 1992, pp. 74-76), focusing on the placement of stress in sentences, including content versus function words, placement of main stress in sentences, and contrastive stress.

In order to teach segmental rules to the group engaged in IPA-SH, two websites were used (http://soundsofspeech.uiowa.edu/index.html#english & http://www.englishaccentcoach.com/complete_consonants.aspx).

For pre- and post-tests, only a sentence-reading task was chosen and four loaded- sentences created by Saito (2011) were adopted. They consisted of 50 words, and 41 among them included pronunciation features that are difficult for Japanese EFL learners (æ, f, v, θ, ð, w, l, ɹ). In this study, a picture description task was not appropriate considering the participants' relatively low speaking skills; Saito and Saito (2016) assert that picture description tasks are not best suited to assessing non-advanced learners' speaking skills. It is likely they would be unable to produce enough utterances to measure, or if they could, their grammatical inaccuracies would impede a fair comparison between the participants.

Table 1
Raters' Background Information

Gender	Nationality	Age	Time spent living in Japan	English proficiency
Female	Mongolian	20	10 months	Advanced
Female	Malaysian	19	4 months	Advanced
Male	Kenyan	25	1 year and 4 months	Advanced
Female	Japanese	21	21 years	Advanced
Female	Ukraine	21	10 Months	Advanced

Procedure

This experiment adopted a quasi-experimental design. To eliminate the potential influence of being taught by different instructors, the same instructor taught the two groups. Prior to the experiments, the author worked with two university students majoring in English education who have advanced English skills in order to examine the suitability of the task for the main experiment.

After confirming the procedure and the difficulty level of the materials through the pilot study with the two students, the author conducted the pre-test prior to the 15 lessons. Each participant read the four loaded sentences (see Materials) in a small room and the performance was recorded with an IC-recorder. Then, the H-SH group received the passage that explains the supra-segmental rules (see Materials) to study on their own, followed by an instructor's follow-up summary of the contents with a Q and A. For the IPA-SH group, the instructor taught the eight difficult segmental features (æ, f, v, θ, ð, w, l, ɹ) using two websites (see Materials), allowing the participants to understand how each sound is produced and how it should sound. A single explanation of the difficult segmental features was not enough for the participants to fully understand the sound system, so in each lesson one feature among the eight was chosen for fuller explanation throughout the 15 lessons. For example, on Day 2, /æ/ was taught, and on Day 3, /f/, covering the eight important segmental features (æ, f, v, θ, ð, w, l, ɹ). The procedure of the experiment is summarized in Table 2.

Table 2
Procedure of Daily Shadowing Activities

Step	H-SH	IPA-SH
1	Listen	
2	Check suprasegmental features individually	IPA introduction
3	Check stressed words with instructor	Transcribe the passage in IPA
4	Listen and practice reading	
5	Silent H- SH (without vocalization)	Silent IPA-SH (without vocalization)
6	H-SH	IPA-SH
7	Script check	
8	H-SH twice	IPA-SH twice
9	Script check	
10	H-SH record	IPA- SH record

A total of 15 lessons were given to the participants. In each lesson, the two groups worked on learning the vocabulary and contents of the 450-550 word passage, which took approximately 30 minutes. Then, the H-SH group and the IPA-SH group engaged in each task as shown in Table 2, working on the selected 40- 60 word passage for about 30 minutes. For the remainder of the class, the participants worked on wrap-up activities. The procedure was adapted from the one proposed in Hamada (2016a) for a shadowing procedure model for listening.

Since some of the tasks in Table 2 may be confusing, further explanation is provided below. In step 3 of H-SH, the instructor checked the stress with the participants, while the participants in IPA-SH transcribed the passage in IPA. In step 5 (Silent H-SH), the learners worked on H-SH without vocalization. Since they did not need to expend their attention on vocalization, they could focus more attention on the haptic gestures. There are two versions of H-SH: a weaker version (with a written script) for less advanced learners, and a stronger version (without a written script) for advanced learners. To avoid cognitive

overload, less advanced learners should shadow with a written script, and add the haptic movement. Shadowing (not with a written script) requires tremendous concentration (Kadota, 2007), so additional physical movement may be difficult for less advanced learners to cope with. With the help of the written script, they will have more cognitive capacity to focus on the haptic movement than they would if they were shadowing with haptic movement without the script. In fact, the pilot study of this technique supports this assumption. On the other hand, the advanced learners have a cognitive capacity large enough to manage to engage in shadowing without a script and with the addition of haptic movement. Considering the proficiency of the participants in this study, the weaker version was chosen.   Therefore, in this experiment, the weaker version of H-SH was adopted. In step 5, learners in the IPA-SH group shadowed the IPA transcription without speaking out loud, attending only to the sounds and the transcription, so, in terms of attention, the silent H-SH and IPA-SH are different from normal listening.

Measurement

In line with the research standard (e.g., Derwing & Munro, 1997), raters' intuitive judgments were adopted for the three categories: comprehensibility, segmental features, and suprasegmental features. These categories were assessed using 1-6 Likert scale. Acoustic measures were not chosen because the research focused on the change in the impression the participants would make on listeners, rather than the detailed acoustic change of their utterances.

Prior to the assessment, a pilot study was conducted. Three proficient non-native speakers volunteered to examine some of the sample speech to assess the potential efficacy of the procedure. In the primary assessment, the raters listened to the participant's recording twice. On the first listening, they rated comprehensibility, and on the second listening, segmental and suprasegmental features. To reduce the likelihood of fatigue, the assessment took place on two separate days, with each session being approximately 90 minutes. After every 5 or 6 recordings, they took a short break.

Before the assessment, the five raters were given training in order to fully understand the procedure, and to ensure they could assess the speech properly. The author explained the three criteria of comprehensibility, segmentals, and suprasegmentals, until they understood each completely; then, the five raters practiced on some of the sample recordings until everybody was ready.

Analysis

Dependent t-tests were performed on each group respectively for the three categories. The alpha level was set at $p < .025$ by way of the Bonferroni correction (i.e., divide .05 by 2). Because it is a pre/post-test design, the magnitude of the pre- and post-progress, r, was calculated. The summary of the results is shown in Tables 3, 4, and 5.

Results

Descriptive Statistics

As the descriptive statistics show, the H-SH group improved on all three of the criteria of comprehensibility, segmentals, and suprasegmentals, while the IPA-SH group also improved on all criteria, but the improvement was less significant than that of the H-SH group (Tables 3, 4, and 5). The plot of comprehensibility displays a clear straight line only for the H-SH group (Figure 1), showing that the participants generally improved to a similar degree. Regarding segmental feature development, only the H-SH group shows a similar degree of improvement relative to comprehensibility. Since the IPA-SH group did not have students with quite low English proficiency at the pre-test, a straight line is not observed, and the results cluster in one area (Figure 2). Regarding suprasegmental features, the IPA-SH

group shows a clear and stable straight line, while the H-SH group shows a rather steep line, meaning the H-SH group's degree of improvement was higher (Figure 3).

Inferential statistics

The t-tests show that the H-SH group statistically improved on all the three features (p<.025), while the IPA-SH group only improved on comprehensibility and segmental features (p<.025) (Tables 3, 4, and 5). In the H-SH group, the effect size (r) is large for comprehensibility and segmental features, and medium for suprasegmental features. In the IPA-SH group, it is only segmental features that show a large effect size, while comprehensibility and suprasegmentals have a medium effect size.

Table 3
Statistics of Comprehensibility for H-SH and IPA-SH groups

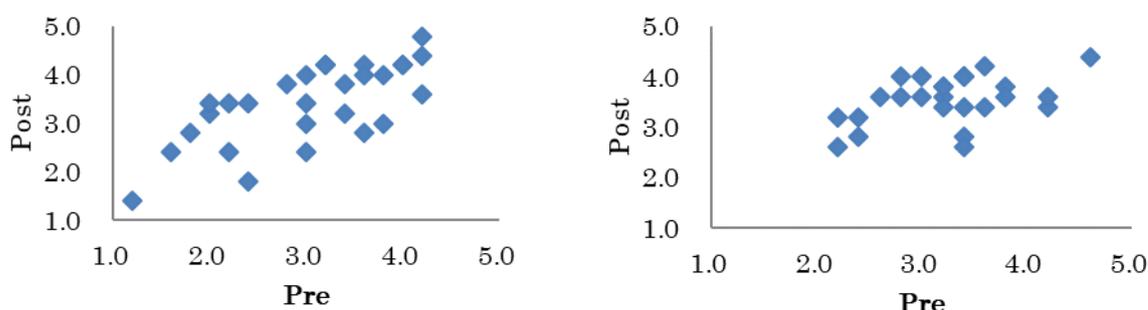|	Pre	SD	Post	SD	t-value	p	r
Haptic	3.03	.855	3.41	.823	-3.10	.005	0.51(L)
IPA	3.24	.585	3.53	.466	-2.68	.012	0.46 (M)

Figure 1. Scattered plots of H-SH (left) and IPA-SH (right) -groups for comprehensibility

Table 4
Statistics of Segmental Feature for H-SH and IPA-SH Groups

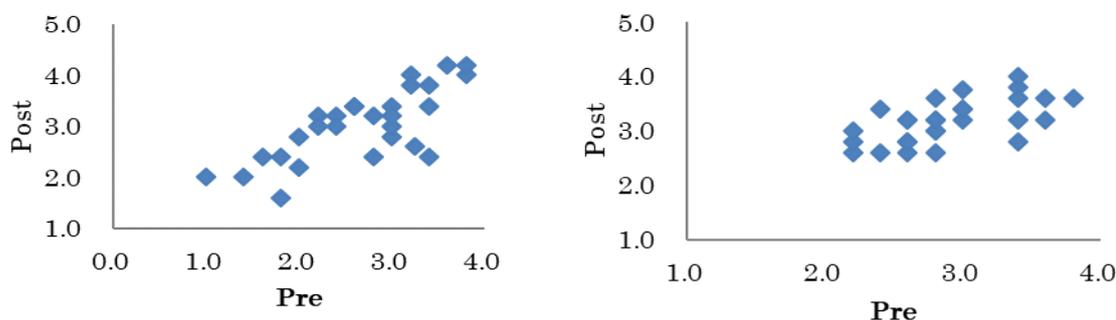|	Pre	SD	Post	SD	t-value	p	r
Haptic	2.67	.745	3.04	.699	-3.83	.001	0.59 (L)
IPA	2.90	.454	3.19	.400	-3.95	.001	0.61 (L)

Figure 2. Scattered plots of H-SH (left) and IPA-SH (right) groups for segmental features

Table 5
Statistics of Supra-segmental Feature for H-SH and IPA-SH Groups

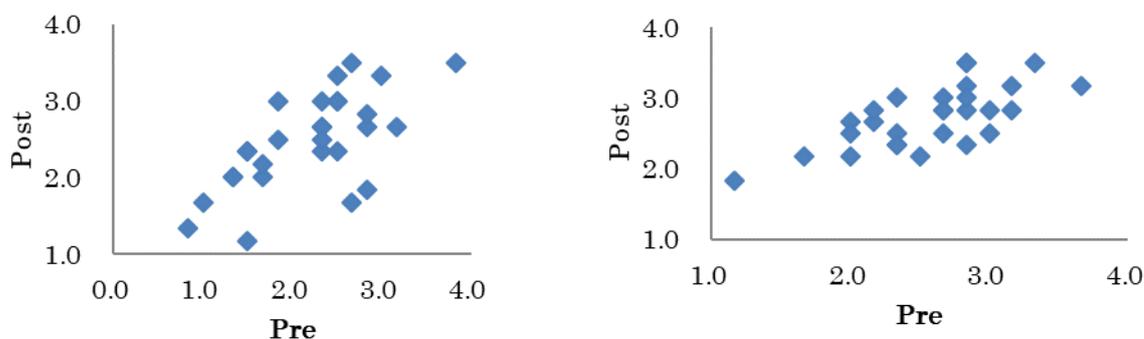|	Pre	SD	Post	SD	t-value	p	r
Haptic	2.19	.707	2.48	.621	-2.79	.010	0.47(M)
IPA	2.57	.528	2.74	.437	-2.21	.036	0.39(M)

Figure 3. Scattered plots of H-SH (left) and IPA-SH (right) groups for supra-segmental features

Discussion

The first RQ looked at whether learners will improve their comprehensibility, segmental productions, and suprasegmental productions by practicing H-SH, and whether learners will improve their comprehensibility, segmental productions, and suprasegmental productions by practicing IPA-SH. Judging from the descriptive statistics and inferential statistics, the H-SH group statistically improved on all the three features, while the IPA-SH group only improved on the comprehensibility and segmental features. In this section, segmental and suprasegmental features are first discussed, then comprehensibility is discussed.

In theory, the IPA-SH group should improve segmental features, and the H-SH group should improve suprasegmental features, because the primary focus of IPA-SH is on segmental features, while the focus of H-SH is on suprasegmental features. In standard shadowing, learners exclusively train their attention on the incoming sounds they hear when shadowing (Hamada, 2016a; Kadota, 2007). In contrast, in this study, their attention was deliberately controlled: in H-SH shadowing, their attention was split between the incoming sounds and stress by adding physical movement, while in IPA-SH, their attention was separated between the incoming sounds and multiple segmental features with the help of the phonetic alphabets. Therefore, the results align with the theory.

It was a surprise that the H-SH group also improved segmental features, with a large effect size (r=.59). Careful interpretation of the plot (Figure 2) shows that the improvement of the lower-level participants contributed significantly to this result. The H-SH group had a higher number of less skilled learners than the IPA-SH group, and the lower-level learners in the H-SH made greater progress. One possible explanation can arise from considering the theory of TPR. TPR bases its theory on a child's initial language acquisition, in which the child does a lot of listening before speaking: the child's listening involves physical responses such as moving, grabbing, and looking, and TPR is designed to be stress-free (Brown, 2001, referring to Asher, 1972). As with this theory, H-SH also involves physical movements and reduced anxiety, and these factors may have improved the learning of the H-SH group, especially that of the lower-level learners. There is also a chance that the intermediate-level learners plateaued, making their progress slow down. In contrast, the learners in the IPA-SH group were directed cognitively to attend only to segmental features, so theoretically they would show improvement in the segmental features.

Taken together, the data in this study suggest the following: the IPA-SH method was helpful for

improving segmentals; the H-SH method was helpful for improving suprasegmentals; and the H-SH appears to help lower-level learners improve their segmentals as well. Still, a further study is necessary, because of a sample size and the exploratory nature of the study.

Next, research has shown that both segmental features and suprasegmental features account for comprehensibility (e.g., Kang, Rubin, & Pickering, 2010; Saito, 2011), so in theory either segmental feature improvement or supra-segmental feature improvement will lead to comprehensibility improvement. The results of this study generally support this theory and it is assumed that segmental improvement boosted the IPA-SH group's comprehensibility, and suprasegmental improvement boosted the H-SH group's comprehensibility.

Limitations

The results have several limitations to be acknowledged. First, the proficiency levels of the two groups were slightly different, in that there were more learners with low-proficiency levels in the H-SH group. Though this was inevitable because of the nature of quasi-experimental design, a future experiment using more balanced classes would make the findings more convincing. In addition, designing a similar experiment including a sample of advanced learners would enable us to use a picture-description task as pre- and post-tests, which may yield further interesting results. Moreover, ideally, a control group would have been set to make the findings more convincing, but this was compromised because of the one-month quasi-experimental design. Second, although the inferential statistics (t-test and effect size) assure the positive effectiveness of both activities, the question still remains of whether the degree of improvement of both groups was satisfactory for the learners. Third, the specification of causes (i.e., the dependent variable) should be carefully interpreted. Both activities consist of several steps in the procedure, so specifying which step is contributing the most to the impact is difficult. Lastly, the criteria used in this study are rather vague because segmental and suprasegmental features consist of several other sub-features. Especially, with H-SH, stress was focused on, and it appears to have led to suprasegmental feature improvement. Suprasegmental features consist of multiple factors, so more specification and attention to intonation and rhythm too would lead to further development of the findings. Lastly, the use of a larger number of raters would make the findings more convincing.

Study 2

Purpose

Study 1 approached the two techniques in terms of their effectiveness on pronunciation development, and study 2 approaches them in terms of the learners' perception of the techniques. This is an attempt to make the study more applicable to classroom teaching by using the Semantic Differential method and a factor analysis. In the context of trying to provide effective teaching techniques, understanding of learners' thoughts and feelings about the methods, as well as their impact on language skill acquisition will help teachers use those techniques more effectively.

Methods

Participants

To collect a data sample large enough for a factor analysis, the data of a new group were added to each group in study 2, in addition to the two groups in study 1. For the H-SH group, data from a group of first-

year Japanese university students who majored in engineering were added. For the IPA-SH group, data from a group of Japanese first-year university students who majored in health science were added. In total, data from 75 participants were collected for the H-SH, and 74 for the IPA-SH. The overall English skill levels of both groups ranged from basic to intermediate, based on the results of a placement test they took at the beginning of the academic year. The first year H-SH group received a haptic-shadowing-based lesson and the first year IPA-SH group received an IPA-shadowing-based lesson. Both groups followed a similar procedure to the ones used in study 1. The main differences between the study 1 procedure were that the groups in study 2 received 8 lessons, while those in study 1 received 15, and the students in study 2 used a different textbook. Still, because the purpose of the study 2 was to examine the impressions of H-SH and IPA-SH, 8 lessons involving the methods were considered to be enough to fulfill the purpose, so these differences are neutralized.

Materials

Out of several possible options for gaining insight into learners' perceptions, the SD method was chosen because it would provide details of the learners' perceptions with the maximum effectiveness in the shortest time. In the SD method, the participants are given a list of adjectives that would describe the target (i.e., H-SH or IPA-SH), and they give their intuitive response to how accurately each adjective explains the target concept. In addition, with a view to comparing the perceptions of the two groups, the SD method enables learners' feelings and perceptions to be quantified. Also, the experiments were conducted in class, and the time available to complete the questionnaire survey was quite limited. For these reasons, the SD method was used.

For the SD method, the same 26-item questionnaire used in Hamada (2011) was adopted for the following reasons. First, this questionnaire was the only one to date that was designed specifically to use the SD method to examine learners' impressions and attitudes toward shadowing. The items are reliable, having been carefully selected through a literature review and in consultation with multiple English teachers at a university and a senior high school (more details appear in Hamada, 2011). In addition, the use of the same questionnaire makes it possible to compare the results of this study with the prior one. Because learners often have negative impressions of standard shadowing, examining how different the perceptions of H-SH and IPA-SH are from standard shadowing would be of value. The questionnaire consisted of 26 adjectives intended to reflect learners' impressions of shadowing (see Results). The participants chose a value from 1-6 on the Likert scale for each pair of adjectives. For example, in the first line, they saw a pair of two adjectives, one positive and the other negative. They had the option to pick greatly positive, somewhat positive, a little positive, a little negative, somewhat negative, or greatly negative. All the items were presented in Japanese.

Procedure

The 26-item questionnaire was distributed to the groups after all the trainings had finished. The instructor explained until everybody understood the procedure giving examples. It took 5-10 minutes. All the participants agreed to the use of the data on condition of anonymity.

Analysis

The data of the two groups were analyzed respectively following the same procedure. An exploratory factor analysis (EFA) was conducted on each group, under the assumption that there would be correlations between the factors. Based on Brown's (2009) five suggestions, four factors were decided: (1) Kaiser's stopping rule, (2) the Scree test, (3) the number of non-trivial factors, (4) a priori criterion, and (5) the percentage of cumulative variance. The Scree plots suggested two factors for the H-SH group and three for the IPA-SH group, but the initial calculation of cumulative variance indicated only 39.81% and

45.54% respectively, values that were too small. Therefore, to ensure that the percentage of cumulative variance exceeded at least 50%, four factors were chosen for each group. This modification made the cumulative variance of the H-SH group 52.75% and that of the IPA-SH group 51.83%. For the interpretations of factor loadings, the criterion of .35 or above was set. After the EFA, 23 items remained in the H-SH group and 20 in the IPA-SH group (see more details in Table 6 for the H-SH group, and in Table 8 for the IPA-SH group).

Results

The EFAs extracted four factors for each group. Briefly, the correlations between the four factors were stronger in the H-SH group and weaker in the IPA-SH group (Tables 7 and 9). While factor 1 of the H-SH group accounts for 26.98 %, and factor 2 accounts for 8.63%, factor 1 of the IPA-SH group accounts for 22.57%, and factor 2 accounts for 12.12%. Moreover, the components of each factor are quite different: Factor 1 of the H-SH group largely consists of combinations of factors 2 and 3 from the IPA-SH group; the items in factors 2 and 4 of the H-SH group are different from those of the IPA-SH group; and factor 3 of the H-SH group largely shares the items of factor 1 of the IPA-SH group. The results for H-SH show similarities with the results of Hamada (2011), but in essence, the learners' perceptions of both H-SH and IPA-SH are different from Hamada's results for normal shadowing.

Table 6
Results of the Factor Analysis of H-SH

item	F1	F2	F3	F4	communality
favorite	.867	.042	-.167	-.080	.700
fun	.752	.131	-.039	-.002	.657
interesting	.720	.021	.124	.072	.619
familiar	.653	-.127	.010	-.180	.429
pleasant	.606	.078	.130	-.182	.622
useful	.538	.187	.177	.018	.580
light	.481	-.223	.144	.223	.235
stable	.400	.106	-.134	.215	.188
deep	-.069	.650	-.035	-.146	.416
clear	.219	.650	-.128	-.049	.546
bright	.154	.567	.085	.102	.478
active	-.063	.557	.061	.183	.298
supreme	-.020	.557	-.002	-.066	.317
severe	-.263	.053	.704	-.236	.542
kinetic	.072	.079	.666	.043	.540
lively	.201	-.046	.619	.131	.472
new	.238	-.201	.549	-.064	.392
loud	-.113	.306	.399	.092	.282
fast	.123	-.113	-.157	-.885	.738
relaxed	-.085	.063	-.158	.615	.472
silent	.046	-.098	-.210	.534	.424
cool	.209	-.040	.158	.410	.191
simple	.100	-.158	-.240	.375	.292

* reversed items are back-worded

Table 7
Correlations Between the Four Factors

Factor correlation matrix	F1	F2	F3	F4
F1	-	-	-	-
F2	.531			
F3	.485	.483		
F4	-.195	-.217	.290	

Table 8
Results of the Factor Analysis of IPA-SH

item	F1	F2	F3	F4	communality
silent	-.885	.138	.021	-.039	709
severe	.851	-.155	-.224	-.099	654
kinetic	.730	.044	-.068	-.023	538
lively	.682	.090	-.215	.245	573
bright	.627	-.053	.217	.234	541
loud	.439	.037	-.019	-.131	210
relaxed	-.384	-.211	.314	.250	282
interesting	.005	.724	-.150	-.041	462
active	-.024	.685	-.159	.119	477
useful	-.322	.668	-.234	.134	447
satisfying	.077	.580	.119	-.045	422
voluntary	.030	.541	.133	-.062	353
supreme	-.113	.533	.001	-.036	242
fun	.136	.516	.220	-.020	464
light	-.131	-.005	.811	-.229	656
simple	-.297	-.121	.554	.177	345
pleasant	.318	.310	.426	-.144	575
familiar	.168	.325	.403	.202	550
clear	.111	-.035	-.101	.669	448
stable	-.117	.057	.016	.634	431

*reversed items are back-worded

Table 9
Correlations Between the Four Factors

Factor correlation matrix	F1	F2	F3	F4
F1	-	-	-	-
F2	.366			
F3	.222	.325		
F4	.054	.326	.070	

Discussion

The constructions of H-SH and IPA-SH will briefly be discussed, followed by the differences and similarities between H-SH and IPA-SH. Then, both will be compared with standard shadowing.

The H-SH consists of four factors: the items in F1 describe participants' subjective feelings about and their preference of H-SH (e.g., favorite, fun, interesting); the items in F2 express their impressions of what H-SH is like (e.g., deep, active, supreme); the items in F3 are more like rough images of H-SH (e.g., lively, kinetic); and F4 consists of items indicating learners' sensitivity to H-SH (e.g., relaxed, cool). Given that F1 consists of the participants' positive feelings, including fun, pleasant, and familiar, H-SH is considered to reflect the important principles of TPR that learners should enjoy learning a language with less stress.

The IPA-SH consists of four factors: F1 indicates the participants' first impressions, and F2 comprises experience-based feelings. Reflecting the fact that IPA-SH mainly involves analytical processes (i.e., reading IPA transcripts), F1 consists of analytical adjectives regarding IPA-SH (e.g., noisy, kinetic, loud). The items in F2 describe general impressions of the nature of IPA-SH（e.g., interesting, active, useful). The items in F3 show IPA-SH's learner-friendliness and approachability (e.g., light, simple, pleasant). The two items in F4 show the clarity of the IPA-SH procedure (e.g., clear, stable).

When compared, the learners' impressions and images of the two activities are generally different, despite both being forms of shadowing. Looking closely at all of the items, only the H-SH group indicated *favorite, new, deep, fast (slow),* and *cool*, while only the IPA-SH group indicated *satisfying* and *voluntary*. Learners also perceive H-SH to be more unique than IPA-SH. As a side note, only H-SH has *favorite*, which suggests that the participants prefer H-SH. F1 of H-SH mainly consists of positive

adjectives that describe learners' feelings toward H-SH, while F1 of IPA-SH contains items that analyze IPA-SH. Given that the cumulative variance of F1 is 26.98% for H-SH and 22.57% for IPA-SH, we can say that the primary perceptions of H-SH are that it is familiar, useful, and enjoyable, while the main perceptions of IPA-SH are that it is a lively and loud activity.

Next, observing the items carefully, we can see first of all that F3 of the H-SH group and F1 of the IPA-SH group are similarly constructed, and IPA-SH contains all the items in H-SH apart from *new*. Most likely, this difference reflects the fact that IPA-SH generally appears to maintain the original style of shadowing (listening and repeating simultaneously), while H-SH seems new in that shadowing is accompanied by physical movement, which stems from an innovative idea, haptic-teaching. The items of F2 in H-SH are split into multiple factors in IPA-SH (deep, clear, bright, voluntary, supreme), and these adjectives also show positive attitudes toward H-SH. In particular, only H-SH contains *deep*. Since learners engaging in H-SH perform multiple tasks such as listening, gesturing, and shadowing, learners might perceive this complexity as *deep*. Second, only IPA-SH contains *voluntary* and *satisfying* in F2. Compared with H-SH, IPA-SH requires a more voluntary attitude because it involves primarily cognitive action. In reality in the classroom, learners could pretend to be working on IPA-SH, because teachers cannot watch the cognitive processes of the learners, unlike the physical action in H-SH, meaning it requires a more voluntary attitude. Moreover, because of its cognitively demanding nature, it may give learners more feelings of satisfaction.

In terms of comparison, the perceptions of the two activities are different in essence: the only similarity is that a large proportion of F1 of H-SH is shared by F2 and F3 of IPA-SH. The components include items associated with value, such as *fun, interesting, familiar, pleasant, useful*, and *stable*, so the participants clearly place a certain degree of value on both activities.

Third, when compared with the results of Hamada (2011), what is interesting is that unlike the results for standard shadowing, neither H-SH nor IPA-SH has negative items. The results for standard shadowing contain five negative factors in total (*useless, shallow, empty, outdated,* and *unpleasant*). Therefore, learners' impressions of both H-SH and IPA-SH are better than those of standard shadowing.

Observing the contents of the different factors, F1 of H-SH shares six items out of eight with F1 in Hamada (2011). However, while F1 in Hamada indicated that shadowing is learner-friendly, only F1 in this study (for H-SH) contains *fun, pleasant,* and *useful*. This implies that H-SH adds extra values to standard shadowing, generating more positive impressions. In addition, only F3 of H-SH contains *new*, while F3 in Hamada (2011) contains the opposite, *outdated*.

Regarding IPA-SH, F1 shows that learners' impressions of IPA-SH were somewhat different from their impressions of standard shadowing. Four items in F1 of this study belong to the same category as those of F2 in Hamada (2011), but they occupy opposite ends of the spectrum. IPA-SH appears to be *kinetic* and *lively,* while the standard shadowing appears to be *calm* and *unlively*. The descriptive image of IPA-SH is active, while that of standard shadowing is motionless. This indicates that for the learners as well as in theory, IPA-SH looks similar but is actually different, and it is the differences that add value to IPA-SH.

In summary, as it has shifted from shadowing for listening to shadowing for speaking, the shadowing activity has evolved to leave a more positive impression on learners. Additionally, H-SH seems rather fresher and newer than IPA-SH. Hamada (2011) has revealed that shadowing is an effective activity for improving learners' listening skills, but because of its demanding and repetitive nature, learners' feelings about it are not necessarily good. In this regard, H-SH and IPA-SH positively contribute to these issues.

Conclusion

As was hypothesized, H-SH and IPA-SH both helped learners to improve their comprehensibility. As was assumed, suprasegmental features of the H-SH group improved, and segmental features of the IPA-SH group improved. In addition, surprisingly, the segmental features of the H-SH group improved as well. In terms of directed attention, the IPA-SH group was controlled to focus on segmental features, and

the H-SH group was controlled to focus on stress in the suprasegmental features. In terms of learners' perceptions, H-SH and IPA-SH appear to have evolved from standard shadowing in terms of learners' perceptions and images. In general, learners have a more positive and fresh image of H-SH and IPA-SH. Specifically, H-SH is recognized as new, and learners seem to be more engaged with it. Still, the interpretation is based on learners' impressions, so further research is needed to make the findings more convincing.

Acknowledgements

The Author

Yo Hamada is associate professor in Center for promotion of educational research and affairs of Akita University in Japan. His current research interests cover listening and pronunciation. His recent publications include 'Teaching EFL Learners Shadowing for Listening: Developing Learners' Bottom-up Skills' by Routledge.

Email address: yhamada@gipc.akita-u.ac.jp

References

Acton, W., Baker, A. Ann., Burri, M. & Teaman, B. (2013). Preliminaries to haptic-integrated pronunciation instruction. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th pronunciation in second language learning and teaching conference* (pp. 234-244). Ames, IA: Iowa State University.

Asher, J. (1977). *Learning another language through actions: The complete teacher's guidebook.* Los Gatos, CA: Sky Oaks Productions

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.

Baker, A. (2014). Exploring teachers' knowledge of second language pronunciation techniques: teacher cognitions, observed classroom practices, and student perceptions. *TESOL Quarterly, 48*(1), 136-163.

Bovee, N., & Stewart, J. (2009). The utility of shadowing. In A. M. Stoke (Ed.), *JALT 2008 conference proceedings* (pp. 888-900). Tokyo: JALT.

Brooks, R. (2013). *Comparing native and non-native raters of US federal government speaking tests* (Unpublished doctoral dissertation). Georgetown University, Washington, D.C.

Brown, D. (2001). *Teaching by principles: An interactive approach to language pedagogy.* NY: Longman.

Brown, D. J. (2009). Choosing the right number of components or factors in PCA and EFA. *JALT Testing & Evaluation SIG Newsletter, 13* (2), 19-23.

Burri, M., Baker, A., & Acton, W. (2016). Anchoring academic vocabulary with a "hard-hitting" haptic pronunciation teaching technique. In J. Tamara (Ed), *Pronunciation in the classroom*: *The overlooked essential* (pp. 17-26). Alexandria, United States: TESOL Press.

Carey, P. W. (1971). Verbal retention after shadowing and after listening. *Perception and Psychophysics, 9* (1-B), 79-83.

Crystal, D. (2003). *English as a global language.* Edinburgh: Cambridge University Press.

Derwing, T. M. (2012). Pronunciation instruction. *The encyclopedia of Applied Linguistics.* Wiley Online Library. doi: 10.1002/9781405198431.wbeal0968

Derwing, T., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition, 20,* 1-16.

ETS. (2015). Test and Score Data Summary for TOEFL iBT® Tests. Retrieved from http://www.ets.org/s/toefl/pdf/94227_unlweb.pdf

Graddol, D. (2003). The decline of native speaker. In G, Anderman & M, Rogers (Eds.), *Translation today: Trends and perspectives* (pp. 152-167). Bristol, UK: Multilingual Matters

Grant, L. (2014). *Pronunciation myths: Applying second language research to classroom teaching*. Ann Arbor, MI: The University of Michigan Press.

Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201-223.

Hsieh, K. T., Dong, D. H., & Wang, L. Y. (2013). A preliminary study of applying shadowing technique to English intonation instruction. *Taiwan Journal of Linguistics, 11*(29), 43-66.

Hamada, Y. (2016a). *Teaching EFL learners shadowing for listening: Developing learners' bottom-up skills.* New York: Routledge.

Hamada, Y. (2016b). Wait! Is it really shadowing? *The Language Teacher, 40*(1), 14-17.

Hamada, Y. (2016c). Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research, 29*(1), 35-52.

Honna, N., Takeshita, Y., & D'Angelo, J. (2012). *Understanding English across cultures*. Tokyo: Kinseido.

Iino, A. (2014). Shadowing renshuga Eigo speaking ryoku to shadowing no ninshiki ni oyobosu koka [Effects of Shadowing Instruction on English Speaking Performance and Perception of Shadowing]. *HoseiTamaronshu, 30,* 105-121.

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility. *Studies in Second Language Acquisition, 34,* 475-505.

Kadota, S. (2007). *Shadowing to ondoku no kaagaku* [*Science of shadowing, oral reading, and English acquistion*]. Tokyo: Cosmopier Publishing Company.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal, 94*(4), 554-556. DOI: 10.1111/j.1540-4781.2010.01091

Lambert, S. (1988). Information processing among conference interpreters: A test of the depth-of-processing hypothesis. *Meta*, *3*, 377-387.

Lambert, S. (1992). Shadowing. *Meta, 37*(2), 263–273.

Lee, J., & Plonsky (2014). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345-366.

Mori, Y. (2011). Shadowing with oral reading: Effects of combined training on the improvement of Japanese EFL learners' prosody. *Language Education & Technology*, *48*, 1-22.

Negishi, M., Tono, Y., & Fujita, Y. (2012). A progress report on the development of the CEFR-J. *Studies in Language Testing, 36*, 137-157.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching & applied linguistics* (4th ed.). London: Longman (Pearson Education).

Saito, K. (2011). Identifying problematic segmental features to acquire comprehensible pronunciation in EFL settings: The case of Japanese Learners of English. *RELC Journal, 42*(3), 363-378.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, *46*(4), 842-854. doi: 10.1002/tesq.67

Saito, Y., and Saito, K. (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research, 21*(5), 589-608. doi: 10.1177/1362168816643111

Saito, K., & Shintani, N. (2016). Do native speakers of North American and Singapore English differentially perceive second language comprehensibility? *TESOL Quarterly, 50(2),* 421-446.

Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, *38*(4), 439-462. doi:10.1093/applin/amv047

Sypyra-Kzlowaska, J. (2015). *Pronunciation in EFL instruction*. Bristol, UK: Multilingual matters.

Shiki, O., Mori, Y., Kadota, S., & Yoshida, S. (2010). Exploring differences between shadowing and repeating practices: An analysis of reproduction rate and types of reproduced words. *Annual Review of English Language Education in Japan*, *21*, 81-90.

Tamai, K. (1992). Follow-up no chokairyoku kojo ni oyobosu koka oyobi "follow-up" noryoku to chokairyoku no kankei. dai 4 kai "Eiken" kenkyu josei hokoku [The effect of follow-up on listening comprehension]. *STEP Bulletin*, *4*, 48-62.

Tamai, K (1997). Shadowing no koka to chokai process ni okeru ichizuke [The effectiveness of shadowing and its position in the listening process]. *Current English Studies*, *36*, 105-116.

Teaman, B. D., & Acton, W. R. (2013). Haptic (movement and touch for better) pronunciation. In N. Sonda & A. Krause (Eds.), *JALT2012 Conference proceedings* (pp. 402-409). Tokyo: JALT.