THE JOURNAL OF ASIA TEFL Vol.10, No. 3, pp. 1-29, Fall 2013

Effects of Tasks and Awareness on Oral Imitation Accuracy by Japanese EFL Learners

Etsuko Ofuka
Kobe Pharmaceutical University, Japan

Joan E. Gilbert
Kwansei Gakuin University, Japan

This study, which is based on a 12-week classroom experiment, investigates the effects of intervention tasks designed for focused listening and awareness raising on the accuracy of Japanese EFL learners' spoken outputs. The study used the pretest/posttest design with two experimental groups (N = 26 each) and two control groups (N = 36 in total). All participants watched short movie clips of 1 to 2 minutes. The participants of the experimental groups did intervention tasks consisting of dictation and oral repetition. One of the experimental groups was given tasks with the following awareness-raising interventions: cloze-dictation tasks with input enhancement and verbal instructions for English prosody. In order to measure the effects of these tasks, we used the learners' elicited imitation accuracy, prosodic accuracy, and self-reported increased awareness of prosody and unstressed words. The findings reveal that only the participants in the experimental group that did awareness-raising intervention tasks demonstrated significant improvement in spoken outputs; the other experimental group participants, who did dictation and oral repetition tasks, did not show much improvement compared to the control groups' participants. The educational implications of this study indicate that educators could help language learners improve their spoken outputs through awareness-raising interventions.

Keywords: elicited imitation, awareness, dictation, prosodic instruction, spoken outputs

Effects of Tasks and Awareness on Oral Imitation Accuracy by Japanese EFL Learners

Teaching pedagogies that stress communicative language practices often include methods and approaches to improve language learners' spoken proficiency; however, effective spoken communication requires language learners to possess aural comprehension competency to receive and produce related spoken outputs.

Language learners who are learning English as a foreign language (EFL) in local educational settings tend to have limited target language exposure outside of the classroom. Lyczak (1979) found that mere listening exposure to a target language could facilitate spoken production. In Lyczak's experimental study, participants who were exposed to an unfamiliar language (Thai) for less than 1 hour for only 4 days produced better results in Thai production trials than the control groups who were exposed to a different unfamiliar language (Japanese) or music (classical). In a more recent study, from a neuroscience perspective, Fadiga, Craighero, Buccino, and Rizzolatti (2002) investigated the effects of passive listening on the speech motor center in the brain. In their experiment, the participants were asked to listen to words with or without the 'r' sound, which involves a tongue-raising movement. It was found that listening to words with 'r' activated the listener's speech motor center significantly more than listening to words without 'r'. The results imply that there is a close relationship between listening and speech production.

Listening exposure or listening for meaning may be sufficient for casual conversations, but we believe that listening for form is also important to convey and receive a message accurately. Thus, we sought effective approaches for improving EFL learners' accuracy of linguistic features of the target language, English. We adopted an elicited imitation method, which in this study means listening to a model utterance and repeating it as accurately as possible, to help the learners concentrate on listening (Vinther, 2002). Elicited imitation can be done vocally (oral repetition) or in the written form (dictation) and can be carried out in a language class by experienced and inexperienced teachers without special equipment. In this study, we used dictation as the main component of the intervention task followed by oral repetition. Oral repetition was included because it has been shown to improve

EFL learners' articulatory outputs (Ellis & Sinclair, 1996) and also because we wanted the participants to confirm the focused features not only by the written form but also orally.

Dictation tasks have been commonly used for decades among language instructors to facilitate EFL learners' listening, writing, vocabulary, and pronunciation skills (Oller, 1971). According to Underhill (2005), dictation tasks can help learners become aware of sounds they may not hear, like short unstressed syllables of content words (p. 202). Dictation can be thought to be more effective in terms of prosody learning than oral repetition. Kaembach and Schlemmer (2008) found that voiced rehearsal was not as effective for retaining the pitch height of the stimulus as silent or no rehearsal. This implies that oral repetition may impede prosody learning because EFL learners may listen to their voices, which may hinder their ability to concentrate on listening and retaining the model utterance's prosody.

Dictation might be not only effective but also easier than oral repetition for Japanese learners of English. This is because they seem to be more familiar with writing English words than pronouncing them in formal educational settings. Thus, Japanese learners might be able to focus on listening more with a dictation task than an oral repetition task, and more focused listening might help them pay more attention to the target words and prosody.

Dictation as an intervention task can be used for developing language learners' intensive listening skills (Chun, 2010). However, we questioned how much and to what extent EFL learners could notice non-salient features such as unstressed grammatical words by elicited imitation tasks that do not stress noticing features. Schmidt (1990, 2010) claimed that awareness is a requirement for noticing and intake of linguistic features of a target language. In more recent research, the efficacy of second or foreign language instructions used in intervention tasks, which lead to awareness of certain linguistic features, has been reported (Akakura, 2012; DeKeyser, 2005; Ellis, 2005; Norris & Ortega, 2000).

When employing awareness-raising interventions, the specific features of the learners' native language need to be taken into account. The Japanese language does not have unstressed words like English does, so it is especially difficult for Japanese English learners to notice those words when they hear them in English. In

order to raise awareness of unstressed words, we employed dictation tasks with input enhancement. Input enhancement is thought to be effective for drawing learners' attention to form-focused features by increasing the salience of the features (Gasparini, 2004; Sánchez, Pérez, & Gómez, 2010).

Conveying the meaning of a message accurately also requires the appropriate use of prosodic features, which is challenging for foreign language learners. In order to facilitate prosodic performance of Spanish learners of English, Ramírez Verdugo (2006) employed computer-assisted intonation training including explicit information about the English intonation system and visual displays of the pitch movement (or intonation). The results revealed the effectiveness of the training on the participants' prosodic performance and their awareness of the English prosodic and sound systems. Although exactly what component(s) contributed to the improvement of the participants' prosodic accuracy in her study is unclear, the results strongly suggest that interventions for raising learners' awareness of focused prosodic features are effective. Thus, we included verbal instructions on English prosody in our intervention tasks.

In the present study, we used DVD clips as a means to provide authentic input and meaning-oriented activities, which are defined as comprehension and discussion exercises, for the control groups. In addition to what the control groups did, two experimental groups also received intervention tasks. One experimental group received intervention tasks consisting of dictation and oral repetition. The other experimental group received intervention tasks of dictation and oral repetition with awareness-raising components (i.e., input-enhanced dictation, which is defined in this article as added features to increase salience of unstressed grammatical words, and verbal instructions for English prosodic features) with the intention of encouraging learners' noticing of certain grammatical and prosodic features.

Our research questions can be summarized as follows:

1. What are the effects, if any, of an intervention task of dictation and oral repetition, which is designed for focused listening, on the accuracy of the participants' spoken outputs?

2. What are the effects, if any, of awareness-raising intervention tasks (of input-enhanced dictation and oral repetition with verbal prosody instructions) on the accuracy of the participants' spoken outputs?

In order to measure the effects, we used (a) the learners' word-imitation accuracy, (b) prosodic accuracy, and (c) self-reported increased awareness of prosody and unstressed words. In this study, word-imitation accuracy refers to the learners' ability to listen to target sentences and repeat the words in the sentences. Oral imitation was adopted here because the use of oral imitation as a reliable measure of language learners' oral proficiency has been documented in the research (Graham, Lonsdale, Kennington, Johnson, & McGhee, 2008; Henning, 1983; Vinther, 2002), and also because oral imitation could measure even learners' implicit knowledge of specific linguistic features (Ellis, 2008). Prosodic accuracy means the learners' ability to repeat the target sentences with natural intonation and pauses. Since the learners' noticing or/and explicit knowledge of English prosodic features and unstressed words may not reflect on the outputs in a short period of time, their self-reported awareness of those components are investigated.

METHOD

Pre and Post-Assessments

The pre and post-assessments consisted of questionnaires and elicited-imitation recordings.

Pre-questionnaire

The pre-questionnaire was administered before the start of the study to obtain information on the participants' English language background. The questions included: (a) what language(s) were used at home, (b) at what age and how long they had lived in another country, (c) whether or not their family had hosted a

foreign exchange student, and (d) learning experiences of musical instruments and singing. All the questions were in English.

Post-questionnaire

The post-questionnaire was administered at the end of the study to gather data on the participants' reflections. The type of questions used for the post-questionnaire are in Appendix A. Three sets of questions in English and Japanese and the rating scales used are shown in Figure 1. In order to ensure that the participants understood the questions, they were written in English and Japanese. The participants were told to indicate their reply to each question by marking a point on the rating scale.

1. Did you enjoy watching the DVD and doing the tasks?
（DVD を見てのこれらのタスクは楽しんでやれましたか。）

not at all somewhat very much
|.....|.....|

2. How did you find repeating the **expressions** accurately?
（英語表現を正確にリピートする作業はどうでしたか。）

very difficult somewhat very easy
|.....|.....|

3. Would you like to do something like this again?
（このようなタスクをする機会があったら、またやってみたいと思いますか。）

No.....O.K.....Yes
(No, I would not!) (Either way is O.K.) (Yes, I would love to!)

FIGURE 1

Three Sets of Questions and the Rating Scales in the Post-Questionnaire

Elicited-Imitation Recordings

Recording material

The recording material consisted of three example sentences followed by eight sentences consisting of 45 words (55 syllables) shown in Table 1. Each item consisted of a beep, a presentation order, and a model utterance, followed by 4 seconds, in which participants were asked to accurately repeat the model utterance immediately after they heard it.

TABLE 1
Sentences Used for the Pre- and Post-Recordings

Section	Sentences
Practice	1. This is a car.
	2. How should I know?
	3. Thank you very much.
Main	S1: I like American food very much.
	S2: Did you eat lunch yesterday?
	S3: I wanted him to know.
	S4: John is the name of my dog.
	S5: How can I help it?
	S6: So I'll see you tomorrow.
	S7: I went to Cambridge University.
	S8: I wanted him to be an engineer.

The eight sentences in Table 1 were selected based on the following considerations, which reflect the authors' teaching experiences with English language learners at the university level in Japan:

1. The sentences consist of common words, and each one has approximately five words so that the sentences would not be too difficult for intermediate-level learners.
2. The material includes question sentences to examine learners' question intonation.
3. The material includes words in the weak form (e.g., articles, prepositions, pronouns) and words with verb or noun morphemes (e.g., *-ed*, *-s*).

4. The material includes relatively longer sentences (with more than five words) to avoid the ceiling effect.
5. The material includes sentences from the DVD clips to examine the effects of exposure to the sessions.

Model utterances of S1, S2, S3, S4, S5, and S8 in Table 1 were utterances of two male native speakers of American English taken from NII-SRC/ERJ Speech Database (Minematsu et al., 2002). This database was used because (a) the utterances were carefully spoken with clear intonation; (b) they are high-quality recordings; and (c) the database includes utterances of the sentences by more than 10 Japanese college students, so we could later use the typical Japanese rhythm or intonation as a comparative measure for each utterance. Those of S6 and S7 were from the DVD *Night at the Museum* (Twentieth Century Fox Film, 2007) that was used for this study.

Recording procedure

The participants were told that they would hear three sentences in the practice section and then eight sentences. They were asked to listen carefully to each sentence and repeat the sentence exactly as they heard it.

The recordings took place individually outside the classroom in a corridor. The participants sat at a desk with a microphone placed 15 cm away from their mouth. The IC recorder (SONY ICD-UX80) was connected to the microphone and placed at the far end of the desk. The cassette player through which the participants heard the model utterances was placed on the left side of the desk. The instructor set the cassette player for each participant, made sure that he or she understood the procedure, and waited until the first sample sentence was presented and recorded.

INTERVENTION TASKS

Session Procedures

There were two experimental groups (E1 and Ea) and two control groups (CTL1 and CTL2). All four groups of participants had eight sessions over 12 weeks. For the sake of understanding the context, all groups were shown all scenes of the movie, and in addition to the eight sessions, classes in which the participants only watched the movie without a task were inserted.

The participants were exposed to authentic language input (movie clips from the DVD, *Night at the Museum*) and taught by the same instructor, who is a native speaker of American English, at a university in Japan. Class instruction was in English. The experimental groups received about 40 minutes of instruction, and the control groups received 20 to 30 minutes, at the end of each class. There was a difference in the amount of time per session because the control groups did not receive the intervention tasks. The procedures are in Table 2.

TABLE 2
Session Procedures

Step	Procedures
1	Difficult vocabulary and expressions from a movie clip were given before viewing.
2	An explanation of the situation in the movie clip was provided.
3	The movie clip was shown in English with Japanese subtitles.
4	The relationships between the characters in the movie clip were confirmed.
5	Students answered multiple-choice cognitive questions and checked their answers.
6	The same movie clip was shown again without Japanese subtitles.
7	The experimental groups received intervention tasks while the control groups did not.
8	New vocabulary and cultural points were discussed.

After watching a movie clip twice, the experimental groups (E1 and Ea) received a dictation paper with a short conversation in the clip including two or three sentences selected for the task. The instructor repeated each sentence selected for dictation three times during which the participants were asked to write down what they heard on the dictation paper. Then, the answers were put on the chalkboard, and the participants corrected their answers with a red pen. Following the dictation, only the participants in Ea received verbal instructions for raising awareness by the instructor to improve their use of English prosody. Both E1 and Ea groups listened to the instructor's voice and repeated the sentences once. The instructor collected the corrected dictation papers.

Interventions for Raising Awareness

Different dictation methods were used for E1 and Ea. The participants in E1 were asked to write down each word in the target sentences with a few of the following exceptions. The participants were not asked to write people's names, and if the sentence was considered to be too long, some words were given on the dictation paper. On the other hand, the participants in Ea did cloze dictation with input enhancement to increase the salience of unstressed words. A sample of the dictation papers used for both experimental groups is shown in Appendix B.

For raising awareness of unstressed words, the dictation tasks for Ea were carefully prepared with the following considerations: (a) The dictation tasks should not be too difficult for the participants in order to facilitate the feeling of accomplishment; (b) the focus points for the participants should be clear in order to enhance the salience of unstressed words. Based on the above considerations, we decreased the number of words for dictation for Ea by 30% compared to that for E1; the dictation tasks for E1 contained 103 items in 23 sentences, but those for Ea had only 74 items in the same 23 sentences and three additional sentences. Three sentences were added for Ea in order to include five additional unstressed words. As a result, the ratio of unstressed grammatical components to content words for Ea's dictation tasks increased to 66% from 50% for E1's tasks. The words contained in all dictation tasks for E1 and those for Ea are summarized in Table 3.

TABLE 3
Summary of Words on the Dictation Papers for the Experimental Groups

Word category	E1	Ea
Content words	**52 (50.5%)**	**25 (33.8%)**
Unstressed grammatical components		
Function words	**42 (40.8%)**	**32 (43.2%)**
[pronouns]	28 [27.2%]	17 [23.0%]
[prepositions]	11 [10.7%]	10 [13.5%]
[articles]	3 [2.9%]	5 [6.8%]
Verb morpheme and contraction forms (e.g., *want**ed**, **I'll**, what**'s**, did**n't***)	**9 (8.7%)**	**17 (23.0%)**
Total number	**103**	**74**

In addition, we adopted a multiple-choice format (e.g., {I / I'm / I've}) for words with features that are considered difficult for Japanese learners of English (27 items out of 74 items) for Ea. The correct dictation rate (averaged across the participants and the words on the dictation tasks) for Ea was 0.95 and that for E1 was 0.91.

For raising awareness of prosodic accuracy, the instructor gave the Ea group verbal instructions that took less than 5 minutes. The instructions (referred to as *prosodic instructions*) included information and examples about English prosody (e.g., stressed/unstressed words, sentence stress, pauses, and intonation). The objectives of the prosodic instructions were to help the participants become aware of: (a) unstressed words or components that do not exist in their native language, which is Japanese; (b) English prosodic features (e.g., intonation patterns, natural pauses, and stress); and (c) the importance of prosody (e.g., different prosodic features result in a difference in meaning). These prosodic instructions were thought to help language learners retain what they have learned for longer periods of time. The detailed description of the prosodic instructions is in Appendix C.

Participants

All participants were non-English-major university students (either first year or second year; aged between 18 and 21) in Japan. The mean age at which they started learning English at school was 13 years old. The participants' English proficiency level was considered to be *intermediate* based on their scores from an English placement test that was administered before the beginning of the academic year. The summary of groups is shown in Table 4.

TABLE 4
Summary of Groups of Participants and Assigned Intervention Tasks

Group	*n*[a]	Intervention task	Year	Sample size[b]
E1	26	dictation & repetition (once)	2nd	20 (8F+12M)
Ea	26	enhanced-input, cloze dictation & prosodic instructions & repetition (once)	2nd	18 (13F+5M)
CTL1	21	control group	2nd	18 (11F+7M)
CTL2	25	control group	1st	18 (9F+9M)

Note. [a]*n* is the number of students originally enrolled in each group; [b]Sample size is the number of students whose recordings were used for analyses; F is female and M is male.

Both of the experimental groups consisted of second-year students; however, two control groups (CTL1, second year; CTL2, first year) were employed because CTL1's initial level determined by a pretest was slightly higher than that of the experimental groups. The initial English level could influence the outcomes. Thus, another control group (CTL2) was included because their initial level was almost the same level as the experimental groups' levels.

A total of 24 participants' recordings were excluded from all the analyses in the study for the following reasons: (a) participants who experienced living overseas for more than 3 months, (b) those who used languages other than Japanese at home, and (c) those who did not participate in either pre- or post-recording. The numbers of participants used for analyses in each group are also shown in Table 4.

RESULTS

Elicited Imitation Accuracy Rate

Elicited imitation in this study was defined as listening to a model utterance and repeating it as accurately as possible, and we used the term *elicited imitation accuracy* to mean how accurately the participants could repeat each word that they heard.

Calculation of Elicited Imitation Accuracy Rate

Each utterance was scored according to the following procedure:

1. One point was given for each word the participant repeated not only accurately but also in the correct order [e.g., for "This is my dog", This (1 point) is (1 point) dog (1 point) my (0 point)].
2. Instead of one point, 0.5 point was given for each word (a) that was repeated but was slightly difficult for the raters to understand due to the speaker's pronunciation; and (b) that was repeated with a morphemic error (e.g., *-ed*, *-s*).
3. No point was taken off for any phrase or word inserted before or after the utterance, but 0.5 point was taken off for any phrase or word that was inserted between words that were repeated accurately [e.g., for "my dog", my (1 point) favorite (-0.5 point) dog (1 point)].

Two raters scored the utterances. One rater is a speaker of American English and the other is a speaker of Japanese. Both raters had more than ten years of experience teaching English to Japanese learners. The assessment was carried out independently, and the inter-rater reliability score was 0.92. The two raters' assessment scores were later compared, and when the scores of an utterance did not match, the raters listened to it together and decided the final score.

The five sentences (S2, S3, S4, S5, and S8 in Table 1) were used for the assessment. We excluded S1 to avoid the ceiling effect because more than 90% of the participants repeated it accurately in the pre-recordings. We also excluded S6 and S7 because they were taken from the DVD used in the sessions.

As an index for measuring the effects of each intervention task, we calculated the elicited imitation accuracy rate for the five sentences: S2, S3, S4, S5, and S8 (referred to as the overall accuracy rate) as follows: Overall accuracy rate (pre or post) for a participant = (sum of scores for the five sentences in the pre- or post-recordings of the group) / 29; where 29 is the total number of words in the five sentences.

Overall Accuracy Rate for Each Group

We used the difference between each group's overall accuracy rate in the post-recordings and that in the pre-recordings as an index for each group's improvement. Figure 2 shows the means of overall accuracy rate in the pre- and post-recordings for each group.

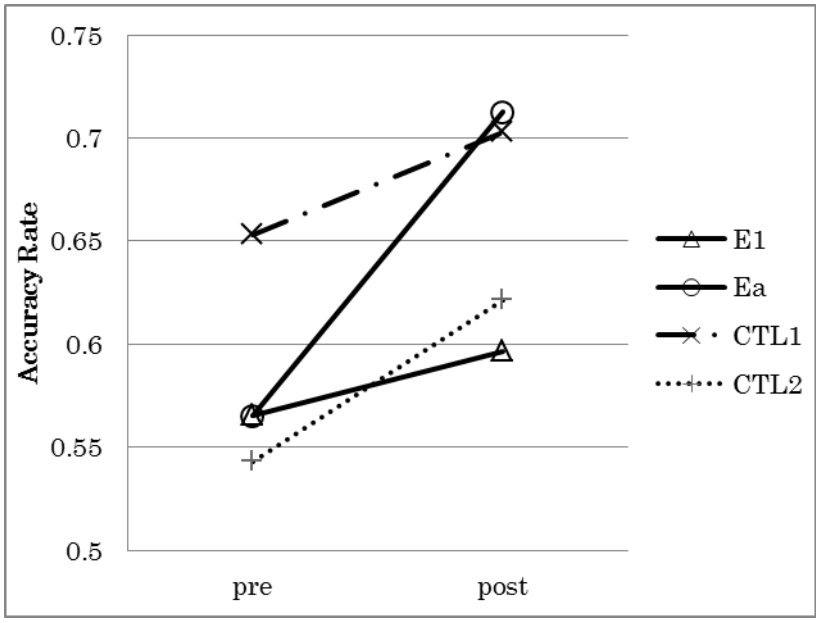

FIGURE 2
Mean Overall Accuracy Rates

A one-way analysis of variance (ANOVA) was performed on the variables of the overall accuracy rate of the post-recordings minus that of the pre-recordings, with the four participating groups. For this (and the subsequent analyses), the alpha level of significance was set to .05, and all analyses were performed using the statistical software package, Excel Toukei 2008, Academic version (SSRI, 2008). This analysis showed a statistically significant group effect, F (3, 70) = 4.57, p = .006. *Post hoc* tests (Tukey's HSD) showed Ea > E1 (p = .005); and Ea > CTL1 (p = .029).

Elicited Imitation Accuracy of Unstressed Grammatical Components

We examined the effects of the dictation tasks and the prosodic instructions on the participants' repetition performance of unstressed, grammatical words and morphemes: Two articles (*the*, *an*); two prepositions (*to* x 3, *of*); five pronouns (*you*, *I* x 2, *my*, *him* x 2, and *it*); and one verb morpheme (-*ed* x 2) were included in the five sentences used for this study (see Table 1). The verb morpheme -*ed* was included only for Ea.

In order to examine each group's improvements, we calculated the elicited imitation accuracy rate averaged across the unstressed components and the participants in each group. Two pronouns, *you* and *I*, were excluded from the analysis in order to avoid the ceiling effect because the highest rates for these words in the pre-recordings were higher than 90%.

Figure 3 shows each group's elicited imitation accuracy rate for the eight unstressed components in the pre- and post-recordings. From Figure 3, we can see that the Ea group showed the highest improvement, but the elicited imitation accuracy rate of the unstressed components for that group was not noticeably different from the other groups. An ANOVA test revealed that there was no significant group effect, F(3, 70) = 1.22, p = .308.

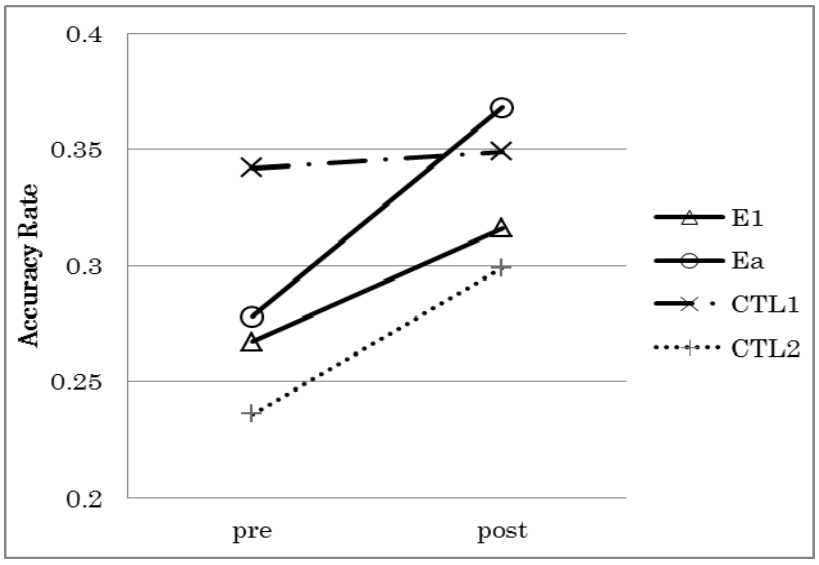

FIGURE 3
Mean Elicited Imitation Accuracy Rates for the Eight Unstressed Grammatical Components

Effects of the Prosodic Instructions

In order to examine how well the participants remembered the sentences practiced in the sessions, we calculated the difference between each participant's elicited imitation accuracy rate for the word *Cambridge* in S7 ("I went to *Cambridge* University") in the post-recordings and that in the pre-recordings. We selected the word *Cambridge* for the following reasons: (a) it was not a familiar word for the participants (based on the elicited imitation accuracy rate for this word in the pre-recordings); (b) it is not a general noun, so the participants were unlikely to hear or use it outside the sessions. This word was included once in two sentences in one session, and all the participants listened to it while they were watching the DVD clip. The participants in E1 then wrote down this word in a following dictation task and corrected it if it was not accurate. The participants in Ea used the

same sentences as E1 but did a cloze-dictation task instead; they did not have to write down the word because it was given on their dictation paper. After the dictation task, they received the prosodic instructions using the two sentences that included this word (see Session 8 in Appendix C).

No group obtained a high elicited imitation accuracy rate ($M \leq .15$) in the pre-recordings. However, the rates in the post-recordings varied from group to group; Ea improved greatly (.00 to .69), whereas E1 demonstrated much less improvement (.00 to .25), and the control groups improved the least (.15 to .25). Non-parametric tests (Kruskal-Wallis) were used for statistical analyses because the values were either 1 or 0. There was a statistically significant group effect, χ^2 (3)= 20.17, $p < .001$. *Post hoc* tests (Scheffé) showed Ea's improvement was significantly greater than those of E1 ($p = .041$), CTL1 ($p = .022$), and CTL2 ($p < .001$).

The smallest improvement for the control groups was expected, but E1's improvement was not impressive regardless of doing the dictation task. On the other hand, Ea outperformed E1 even though the participants in Ea did not write down the word on their dictation paper. This strongly suggests that the prosodic instructions were effective for the participants' memorizing what they learned.

Post-questionnaire

A post-questionnaire was given to the participants at the end of the study to gather data on the participants' reflections of the sessions. The questionnaire included the level of: (a) the enjoyment of the sessions, (b) the meaningful effect of the tasks for listening and oral repetition, and (c) awareness of English prosody gained from the sessions. The questions are in Appendix A.

The replies of the participants in the four groups are shown in Table 5. Almost all participants (98%) reported that they enjoyed the task, and 89% reported that they wanted to do it again. Most participants in both the experimental and control groups reported that the sessions were especially useful for listening; an average of 78% of the participants in the four groups acknowledged the usefulness of the sessions for listening. The majority of the participants (67%) also acknowledged the usefulness

for oral repetition even though they did oral repetition only once after dictation. Mean scores for increased awareness varied from group to group, but Tukey's tests showed no statistically significant difference in pairwise comparison, with the exception of CTL1 vs. CTL2 for the increased awareness of intonation (p = .030).

CTL2 had the highest score for increased awareness of intonation (C-1 in Table 5). CTL2 was the only group consisting of first-year students, and it is speculated that the impact of the lessons by the instructor, who was a native speaker of English, might have been stronger for first-year students than for second-year students. The experimental groups, E1 and Ea, had higher scores (M = 82 and 85.5 respectively) than CTL1 (M = 67.5). Ea's high score was expected because the verbal instructions on prosodic features including intonation were given to only this group. The intensive listening E1 did for the dictation tasks might account for their high score.

TABLE 5
Replies of Participants Who Received the Post-questionnaire

Question category	Ea (n=18)	E1 (n=18)	CTL1 (n=15)	CTL2 (n=17)
(A) Enjoyment	[0 (*not at all)* — 100 (*very much*)]			
(A-1) Enjoyed?	97	97	96	**100**
(A-2) Want to do it again?	86	83.5	86	**100**
(B) Are the tasks useful to …?	[0 (*not at all)* — 100 (*very much*)]			
(B-1) Listening	**83.5**	75.5	78.5	73.5
(B-2) Oral imitation	**69.5**	64	67.5	67.5
(C) Increased awareness of …?	[0 (*not at all)* — 100 (*very much*)]			
(C-1) Intonation	85.5	82	67.5	**94**
(C-2) Unstressed words	**74**	62	64.5	67.5
(C-3) "Real" English sound	68	58	**84.5**	73.5

Note. The highest scores among the four groups are in boldface, and the lowest scores are underlined.

For increased awareness of unstressed words (C-2 in Table 5), Ea had the highest score. Ea also showed the most improvement in the elicited imitation accuracy rate for unstressed words, although the difference was not statistically significant. The dictation method for Ea with the multiple-choice format used for unstressed words and/or the prosodic instructions might have contributed to this group's progress. On the other hand, E1 had the lowest score. This group did the dictation including unstressed words, but it is speculated that there might have been too much stimulation which resulted in distracting attention from the target unstressed words.

For awareness of the "real" English sound (C-3 in Table 5), the control groups had higher scores than the experimental groups. The control groups watched DVD clips, which included authentic input through the conversations, whereas the experimental groups listened to the instructor's enunciation for the dictation tasks after watching the scene. This instructor's enunciation might have influenced the participants' perception of English sounds.

In summary, the participants in Ea, who received the awareness-raising interventions, showed higher scores than E1 for all of the above post-questionnaire items, suggesting the positive effects of this type of interventions. The E1 group, on the other hand, had the lowest scores for the increased awareness of unstressed words and that of the "real" English sound. This may suggest that only short dictation tasks (like the ones used in this study) might not have much impact on raising language learners' awareness.

Prosody

We performed a preliminary prosody assessment because many participants reported in the post-questionnaire that they increased awareness of prosody [see (C) in Table 5]. We examined the question intonation of the participants, who repeated the S2 utterance ("Did you eat lunch yesterday?") accurately in both pre- and post-recordings. Two raters (two native speakers, one male and one female, of American English with an experience of teaching English in Japan for over 15 years) listened to the utterances, which were presented in random order. Each stimulus presented to the raters consisted of a beep, a presentation order, a model

utterance, and an utterance to be assessed followed by 4 seconds, during which assessment was carried out. The raters listened to all of the utterances once while assessing each utterance in terms of how accurately the utterance imitated the model utterance. Then they listened again to the same utterances, circling any factor(s) that they thought might be influencing their assessment of the imitation. The factors were: (a) intonation (*too flat* / *not appropriate sentence-initial or sentence-final intonation*), (b) pause (*long* / *frequent*), (c) stress/rhythm (*wrong stress* / *shotgun rhythm* / *smoothness*), and (d) speed (*too slow*).

In order to examine any prosodic improvement, we first counted the number of circles for each factor in each group's pre- and post-recordings. Then we averaged the number of circles across the participants in each group (Mean C) for each rater. Then we calculated an index for progress by subtracting MeanC in the post-recording from that in the pre-recording. Plus values mean the number of participants whose utterance improved in terms of question intonation in the post-recording.

Table 6 shows the results averaged across the two raters for the two factors, flatness and sentence-final intonation, which were found to be a problem for the majority of the utterances in the pre-recording. We can see from Table 6 that the Ea group showed noticeable improvement in intonation range. It should be noted, however, more of the utterances of this group in the pre-recording was assessed as having problems with this factor compared to the E1 and control groups, and Ea's performance as a group in the post-recording did not even reach that of the control groups. However, compared to the performance of E1, we can conclude that Ea's improvement was substantial. For the sentence-final intonation, although Ea's performance as a group was the worst in the pre-recording, that in the post-recording was assessed as the best. Although the data are limited, it can be concluded that Ea showed superior improvement in the sentence-final intonation for a question sentence.

TABLE 6
Prosody Assessment of the Participants' Utterances on Sentence Intonation

Group	flatness			sentence final intonation		
	Pre	Post	progress (pre-post)	Pre	Post	progress (pre-post)
Ea (*n*=7)	.43	.29	**.14**	.43	.07	**.36**
E1 (*n*=6)	.25	.33	-.08	.13	.42	-.29
CTL1+CTL2 (*n*=7)	.14	.14	0	.21	.14	.07

Note. Two raters circled any factor(s) that they thought might be influencing their assessment of the participants' utterances. The number of circles was counted for each factor in each group's pre- and post-recordings and was averaged across the participants and the two raters. The highest scores in the *progress* row are in boldface and the lowest scores are underlined.

DISCUSSION AND CONCLUSION

Our first research question was: What are the effects, if any, of the intervention task of dictation and oral repetition for focused listening on the accuracy of the participants' spoken outputs? The positive effects of the intervention task were not confirmed from the results in this study; the experimental group which did this task, E1, did not show much improvement in terms of overall accuracy rate, elicited imitation accuracy rate for unstressed words, and increased awareness of English prosody, compared to the control groups.

The effectiveness of this intervention task was not conclusive in this study, which might be related to the amount or/and method of dictation. The results differed from Chun's (2010), which reported that dictation of rapid speech did have positive effects on the teaching English as a second language (TESL) students' listening/speaking ability. The main difference was that the amount of time spent on the dictation tasks was far greater for Chun's participants than for ours. Another factor that might have influenced the results is the number of oral repetitions after the dictation; the repetition was done only once, which might not have been enough for the participants to internalize the pattern. In addition to the above factors,

learners' productive skills could be delayed compared to their declarative knowledge. There could be long-term effects of dictation, which may be worthy of further investigation.

The second research question was: What are the effects, if any, of the awareness-raising intervention task on the accuracy of the participants' spoken outputs? In order to examine the effects of raising awareness, we provided the other experimental group, Ea, with the following: (a) input-enhanced dictation for increasing the salience of unstressed grammatical words, and (b) a less-than-5 minute verbal instruction on English prosody after each dictation task per session. The results confirmed the effectiveness of the sessions for Ea. Compared to the other participating groups, Ea showed superior progress in overall accuracy. However, compared to this drastic improvement, Ea showed only modest improvement in elicited imitation accuracy for the unstressed grammatical words/morpheme, although this group's progress was the largest among all the participating groups. With these results, the effectiveness of this dictation method which highlights unstressed grammatical words was useful in the present study but not completely conclusive for generalizing the results. The amount or/and frequency of the dictation may not have been enough for raising the participants' awareness to the point that they could make observable improvements. In addition to dictation, explicit explanation related to the use of the grammatical word/morpheme might have contributed to improved outcomes.

The Ea group, which received awareness-raising intervention tasks, showed noticeable progress in increased awareness of English intonation. Although the data were limited, the improvement was confirmed not only in awareness but also in intonation accuracy in terms of intonation range and sentence-final intonation of a question sentence. Compared to Ea, CTL2 showed the highest score in increased awareness of intonation, but CTL2's raised awareness did not directly lead to improvement in their production of intonation. We can speculate from the results that the prosodic instructions, which emphasized the meaning difference associated with prosodic changes, helped the EFL learners in this study bridge the gap between increased awareness and actual improvement of intonation production.

It should be noted that, after eight sessions spread over 3 months, Ea showed noticeable improvement in question intonation but just modest improvement in elicited imitation accuracy for unstressed words. The participants in this study were native speakers of Japanese (classified as a pitch-accent language). As a result, the participants might have been more sensitive to aspects of intonation than those of duration, and changing their intonation might have been easier for them than changing durational features. Japanese EFL learners' increased sensitivity to aspects of intonation over those of duration was found in a study on sentence-stress perception in English (Ofuka, Mori, Gilbert & Kiritani, 2009). This tendency may have resulted in quick improvement in intonation but not in recognition and imitation of unstressed words in the present study.

The verbal instructions about English prosody, which the participants in Ea received, seemed to have a great impact on the participants, contributing to the retention of the practiced utterances in the session. The participants in Ea showed significant improvement in the elicited imitation accuracy for the word, *Cambridge*, which was included in two of the selected sentences for dictation in a session, compared to both E1 and control groups. The participants in E1 wrote down the word and corrected it, if necessary, but their improvement on the elicited imitation accuracy for the word was only slightly better than that of the control groups. On the other hand, the participants in Ea did not write down the word for dictation but received the prosodic instruction for the sentences that included the word.

In summary, significant progress was confirmed for the participants in Ea in overall accuracy. Noticeable improvements were also seen in the elicited imitation accuracy for unstressed words, the self-reported increased awareness of intonation, and prosodic accuracy for the intonation of a question utterance, but the results did not reach a significance level of .05. It was also shown that the prosodic instructions, which were included as part of the awareness-raising interventions, contributed to the retention of the practiced sentences. These results confirmed the effectiveness of the awareness-raising intervention tasks, especially the prosodic instructions, in the present study. The prosodic instructions took less than 5 minutes per session and did not require special equipment. Thus, this type of instruction could be implemented in any language lesson easily. The educational implications of this study indicate

that educators could help language learners improve their spoken outputs through awareness-raising interventions. What type of instruction is effective may depend on various factors including learners' first language and the target linguistic features. The instructions in this study focused on prosodic features. Even though language learners, for example, produce questions with a rising intonation, their level and/or manner of the final pitch movement could make their questions be perceived as statements instead. These subtle features may be very difficult for language learners to notice even if they listen attentively. Thus, the learners can benefit from awareness-raising interventions such as explicit explanation about the focused features and associated meaning difference.

Finally, the results of this study indicated that raising awareness had a positive influence. However, there are several limitations. The target grammatical and prosodic features included in the pre and post tests were limited in variety and levels of difficulty. In addition, this study was conducted in a classroom environment with a convenience sample, so the findings could have been influenced by other factors, such as the participants' gender, amount of exposure to English outside the classroom, varying learning styles, and motivation. Further studies considering the above factors are recommended in order to generalize the results of this study.

THE AUTHORS

Etsuko Ofuka has an M.A. (Japanese language pedagogy) from the University of Toronto and a Ph.D. in Psychology specializing in speech science from the University of Leeds. She is an EFL Instructor at two universities in Japan. Her research interests include experimental phonetics, speech perception, and speech production.
Email: ofuka@st.kobepharma-u.ac.jp

Joan E. Gilbert has an M.A. (TESOL) from Columbia University, Teachers College and is currently working on her Ph.D. in Education, specializing in global

and comparative education. She is an EFL Instructor at Kwansei Gakuin University in Japan. Her research interests include awareness training in addition to listening and communication.
Email: z98325@kwansei.ac.jp

REFERENCES

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research, 16*(1), 9-37. doi:10.1177/1362168811423339

Chun, L. S. (2010). Developing intensive listening skills: A case study of the long-term dictation tasks using rapid speech, *The English Teacher, XXXIX*, 94-120.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55*, 1-25. doi: 10.1111/j.0023-8333.2005.00294.x

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49A*(1), 234-250. doi:10.1080/713755604

Ellis, R. (2005). Principles of instructed language learning. *Asian EFL Journal, 7*(3), 9-24.

Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics, 18*(1), 4-22. doi: 10.1111/j.1473-4192.2008.00184.x

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience, 15*, 399-402. doi:10.1046/j.0953-816x.2001.01874.x

Gasparini, S. (2004). Implicit versus explicit learning: Some implications for L2 teaching. *European Journal of Psychology of Education, XIX*(2), 203-219. doi: 10.1007/ BF03173232

Graham, C. R., Lonsdale, D., Kennington, C., Johnson, A., & McGhee, J. (2008). Elicited imitation as an oral proficiency measure with ASR scoring. Retrieved May 7, 2012 from http://www.lrec-conf.org/proceedings/lrec2008/index.html.

Henning, G. (1983). Oral proficiency testing: Comparative validities of interview, imitation, and completion methods. *Language Learning, 33*(3), 315-332. doi: 10.1111/j.1467-1770.1983. tb00544.x

Kaembach, C. & Schlemmer, K. (2008). The decay of pitch memory during rehearsal (L). *Journal of the Acoustical Society of America, 123*(4), 1846-1849. doi:10.1121/1.2875365

Lyczak, R. A. (1979). The effects of exposure to a language on subsequent learning. *Language and Speech 22*(1), 81-88. doi: 10.1177/002383097902200106

Minematsu, N., Tomiyama, Y., Yoshimoto, K., Shimizu, K., Nakagawa, S., Dantsuji, M., & Makino, S. (2002). English speech database read by Japanese learners for CALL system development. Retrieved May 7, 2012 from http://www.lrec-conf. org/proceedings/lrec2002/

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417-528. doi: 10.1111/0023-8333.00136

Ofuka, E., Mori, Y., Gilbert, J. E., & Kiritani, S. (2009). Duration as a cue to perceive English sentence stress by Japanese learners of English vs. native speakers of American English. *Journal of the Phonetic Society of Japan, 13*(3), 90-100.

Oller, J. W., Jr. (1971). Dictation as a device for testing foreign-language proficiency. *ELT Journal, XXV*(3), 254-259. doi:10.1093/elt/XXV.3.254

Ramírez Verdugo, D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness, 15*, 141-159. doi:10.2167/la404.0

Sánchez, R. C., Pérez, A. S., & Gómez, P. C. (2010). An attempt to elaborate a construct to measure the degree of explicitness and implicitness in ELT materials. *International Journal of English Studies, 10*(1), 103-129.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan & I. Walker (Eds.), *Proceedings of CLaSIC 2010* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158. doi:10.1093/applin/11.2.129

SSRI (2008). Excel Toukei [add-in computer program for Microsoft Excel]. Tokyo: Social Survey Research Information, Co.

Twentieth Century Fox Film, 20th Century Fox Home Entertainment Japan (2007). Night at the museum [DVD]. California, U.S.

Underhill, A. (2005). *Sound foundations: Learning and teaching pronunciation*. Oxford, UK: Macmillan.

Vinther, T. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics, 12*(1), 54-73. doi:10.1111/1473-4192.00024

APPENDIX A

Post-Questionnaire Questions

Question Category[a]	Questions
A-1	Did you enjoy watching the DVD and doing the tasks?
A-2	Would you like to do something like this again?
B-1	Did it help you improve your listening?
B-2	Did repeating (saying) the expressions become easier for you?
C-1	Compared to before, did you become more aware of natural English intonation?
C-2	Compared to before, did you become more aware of contractions (e.g., *I'll*, *what's*) and weak forms of words (e.g., *a*, *of*)?
C-3	When you listened to the expressions in natural English, did you find that they sounded different from the English you were used to before?

Note. [a] The question category is described in the text.

APPENDIX B

Dictation Papers in Session 1 for the Experimental Groups, E1 and Ea

(1) Dictation paper for E1

Larry:Hey, Mike!

Mike:Hey, Larry. (_____) (_____) (_____)?

Larry: (_____) (_____) (_____) Nicky?

Mike: (_____) (_____) (_____) /

(_____) (_____) (_____) with Erica.

Mike: It was half day today. Uh, "Parent Career Day".

(2) Dictation paper for Ea

Larry: Hey, Mike!

Mike: Hey, Larry. How (_____) doing?

Larry: Have (_____) seen Nicky?

Mike: { I / I'll / I'm } pretty sure { they / he / she } went home with Erica.

Mike: It was half day today. Uh, "Parent Career Day".

APPENDIX C

Prosodic instructions

Session	Selected sentences	Focused points
1	**How you** doing?	-"how you …" vs. "how **are** you …" (casual vs. non-casual style) - rhythm difference (between English and Japanese) - sentence-final intonation
2	I've got a job interview with …	tone of voice and associated meaning
3	Hey, it's me, so **I** got the job. So **I'll** see you tomorrow.	contraction form (*I'll* vs. *I*)
4	**You'll** have to excuse me though.	contraction form (*you'll* vs. *you*)
5	Can I buy you **a cup of coffee** or something?[a]	- sentence-final intonation - article usage (countable vs. uncountable nouns) - linking sounds ("a cup of")
6	Can I talk **to** you about this, please?[b]	- sentence-final intonation - meaning difference (*to* vs. *with*)
7	**Wh**at's wrong?	- intonation and pauses

	What's going on? **Wh**at are you talking about? **Wh**y didn't you tell me?[c]	for Wh-questions and associated meaning difference
8	I went to Cambridge University. **You** went to Cambridge?[d]	- statement vs. Question intonation - stress on *you* and associated meaning difference

Note.

[a] After the dictation, but before repeating, the instructor gave an example of how the sentence would sound if there were no pauses and if the intonation at the end was not raised.

[b] After the dictation, but before repeating, the instructor gave an example of how the intonation would be different without the "please" at the end. Awareness was provided for the sentence-ending visually on the board and verbally.

[c] The focus was on the difference between a "worried" voice and a "regular" voice.

[d] The difference in the intonation just on the final word and the gradual raising were demonstrated.