Effects of Reading Aloud Using a Chant Method: A Comparison of Acoustic Analysis and Human Ears

Kazue KAWAI
Miyagi University

Abstract

The purposes of this study are to investigate how a method of jazz chants can be used as pronunciation practice in Japanese adult EFL classes and to consider what element affects the intelligibility of Japanese EFL learners’ pronunciation by making a comparison between an acoustic analysis and a human ear evaluation. Acoustic data were extracted from a part of Kawai’s (2014) study. In the previous research, 52 participants, classified by their English proficiency level into three groups (high, middle, low), chanted from a text for approximately 15 minutes in their English classes for a period of five weeks. The ability to control the duration of the inter-stress interval (ISI) was used as a criterion for acquiring an English stress-timed rhythm. The present study used 20 participants’ (low English-proficiency group) acoustic data from the above, and two native English speakers evaluated these data using 9-point scale. According to two-way repeated measure ANOVA, the main effect for sentence type and time (pre-and post-test) were statistically significant. Moreover, the correlation between ISI’s duration and the raters’ scores were observed according to each correlation ratio and scatterplot.

1. Introduction

“Jazz chants” as a chant method for teaching and learning English was originally developed by Carolyn Graham in 1978. She was an English as a Second Language (ESL) teacher and jazz musician in New York who provided language students with a rhythmic means for improving speaking and listening skills (Graham, 1978; Richard-Amato, 2003; Segawa, 1995 as cited in Tazaki, 1995). Coincidentally, around the same time, mainstream language teaching and learning underwent a change. Along with the rise of the communicative approach, the goal of pronunciation changed from accuracy to fluency. It is very likely that this tendency promoted a gain in the popularity of jazz chants in ESL classes all over the world. As a consequence, a great number of researchers (Anderson-Hsieh, Johnson & Koehler, 1992; Gilbert, 2005; Hahn, 2004; Kawashima et al., 1999; Kenworthy, 1987; Morley, 1987, 1991; Munro & Derwing, 1995; Pennington, 1987, 1989) began to focus on the importance of teaching suprasegmental aspects, such as stress, rhythm, and intonation, to foster the intelligibility of learners’ pronunciation.
Richard-Amato (2003) stated the popularity of jazz chants: chanting English texts aloud with the beat of a drum helps to reduce anxiety in students and promotes ego permeability in English classrooms. Renner et al. (2017) reexamined Guiora’s (1972) study and mentioned the relationship between ego permeability and learning foreign language again. Moreover, Celce-Murcia, Brinton & Goodwin (1996) pointed out that using jazz chants is a good way to train students to read English texts more fluently, rather than word by word, and that creating individual chants through substitution drills is also helpful for learners (pp. 297–298). Today, the use of jazz chants is a well-accepted methodology in English teaching classrooms across the world.

In Japan, since the publication of Graham’s (1978) Jazz Chants, such chanting has been gaining popularity among teachers of young EFL learners and also many other types of original chants have been made. These chants have been much used as an English language classes in private elementary schools and private English schools, especially with young learners. At present, it is commonly and widely accepted for chanting to be used as an activity in English language activity classes also in public elementary schools in Japan. In general, the expectation is that chants may help young EFL students to become familiar with English stress-timed rhythms, improve their motivation for learning English, and raise their awareness of cross-cultural knowledge.

However, chant methods are not often used to teach Japanese adults, including college students, in English language classes. Furthermore, there have been only a few empirical studies on the effects of chanting in educational settings, and its procedures have not yet been clearly revealed. How does it help to teach and encourage learning English? Can a chant method really help Japanese EFL learners acquire the proper stress-timed rhythm for English as Carolyn Graham (1978) stated? There are many unanswered questions about the function of chanting as a pedagogical tool in English language classrooms.

Therefore, this study focuses on two points: How does a chant method work as an instruction in pronunciation for college students? What elements of Japanese EFL learners’ English pronunciation influence native English speakers’ ratings after receiving instruction in chanting? In this present study, the word “chants” is used as rhythmically repeating words and short phrases aloud while listening to a rhythmic drum beat and to a model reading by a native speaker at the same time; in other words, it represents “parallel reading or shadowing using a jazz chant rhythm.” In addition, the term “adults” includes “college students” in this study.

2. Research Background

In this section, the author gives an overview of the pedagogy of pronunciation in English classrooms, a summary of experimental studies using chants in adult EFL classes, and the purpose of the present study.
2.1 Pronunciation instruction

In general, a pedagogical history of pronunciation is divided into the following three periods: 1940–1950s (audio-lingual approach based on Behaviorism), 1960–1970s (neglect of pronunciation instruction based on Cognitivism), 1980–present (more focus on intelligibility and comprehensibility based on the communicative approach) (Celce-Murcia, et al., 1996; Morley, 1991; Murphy, 2003; Seidlhofer, 2001).

Now we can say it is the time of “World Englishes.” There have still been the issues to be discussed in EFL or ESL classroom settings, for example suprasegmental oriented instruction vs. segmental oriented instruction, “to whom intelligibility” or “to whom comprehensibility.” Derwing and Munro (2015) mentioned various factors in learning and acquiring pronunciation: AOL (age of L2 learning), LOR (length of residence), motivational influences (e.g., L2WTC: Willingness to Communicate), aptitude, and instruction. They also pointed out that the critical period regarding pronunciation is still controversial and also that there is no reason to believe in the unintelligibility of adult EFL learners. They argued the importance of pronunciation instruction, although some other researchers believed that classroom pronunciation instruction was ineffective in the 1970s.

In Japan, in line with the mainstream communicative approach, the Ministry of Education, Culture, Sports, Science and Technology has promoted a more communication-oriented curriculum through its Guidelines (Gakushu shido yoryo) since 1989. However, regardless of the effectiveness of segmental or suprasegmental instruction, pronunciation instruction itself is not yet adequate in junior and senior high schools in Japan (Kawashima et al., 1999). Teshima (2011) suggested that the situation has not changed very much since the 1970s, with most students producing “katakana English” although present students spend more time speaking English in class than before. Sugito (1996) and Sudo (2010) also mentioned that most Japanese EFL learners tend to stress every individual word when they read English texts. In addition, according to their survey, Shibata et al. (2008) reported that although many teachers realized the importance of pronunciation instruction, half of them had less confidence in teaching pronunciation. They also pointed out that the various styles of English, LFC (Lingua Franca Core), or EIL (English as an International Language) have made it difficult to set clear goals in the classroom and have confused English teachers.

Looking back at the above history, three main issues can be highlighted in Japanese EFL classes: less time to teach pronunciation, no clear goal in teaching pronunciation, and lack of teachers’ confidence in teaching pronunciation. However, such a tendency is seen not only in Japan but also in ESL settings across the world. Derwing and Munro (2015) mentioned that English teachers do not have sufficient skills to teach pronunciation and that they have very limited access to professional development in this area (pp. 80–81). Moreover, Szpyra-Kozlowska (2013; 2014) pointed out that the teaching of pronunciation is often neglected. She explains the reasons: pronunciation is regarded as the most difficult aspect of another language to master; most
English teachers tend not to consider pronunciation as important because many international language examinations attach little importance to the examinees’ English pronunciation in comparison to other language skills; there is less space given to pronunciation instruction in materials such as course textbooks; emphasis is placed on fluency rather than accuracy and pronunciation errors are tolerated; many non-native teachers are reluctant and lack confidence in teaching pronunciation because of the poor quality of their own pronunciation; appropriate and effective training courses in the teaching of pronunciation are not conducted enough (Szpyra-Kozlowska, 2014, pp. 4–6).

2.2 Using jazz chants in adult EFL classes

Nakano (1997) conducted research in jazz chant rhythm instruction over a period of 12 weeks. The participants were 10 college students grouped by proficiency level. There were five high-proficiency and five low-proficiency students. She used 16 test sentences, which were arranged by four types: one unstressed syllable in ISI1, two unstressed syllables in ISI2, three unstressed syllables in ISI3, and four unstressed syllables in ISI4. The ISI duration was measured between the voice onset of two stressed vowels. After 12 weeks of instruction, the compressed rate of the low group was significantly higher than that of the high group. As a result, she argued that the instruction was especially effective for low-proficiency learners and that rhythm practice such as jazz chants should be undertaken during the beginning stages of English language instruction.

Kung (2013) conducted a study of 30 participants who were all college or graduate students in an intermediate (TOEIC 550–650) EFL speaking class at a language institute in Taiwan. Over a period of 12 weeks, the participants chanted jazz chants as a warm-up activity for about 20 minutes before their weekly instruction. In pre- and post-tests, the participants took a listening test and a speaking test and completed a questionnaire and an interview. Significant differences were observed in the listening test, the speaking test, and the questionnaire. The study concluded that the use of jazz chants improved EFL learners’ listening and speaking skills, the use of music fostered EFL students’ learning interests and motivation, and the use of jazz chants could give students more confidence in speaking English.

In Kawai (2009), 14 Japanese adult EFL learners, whose English proficiency was Eiken 3–4 on a grade-level test score, were given 11 weeks of chanting rhythm instruction. The participants chanted some sentences while listening to the CD for about 40 minutes in every class. In pre- and post-tests, the participants took a reading test, an interview test, and a listening test. These speech data were evaluated by two native speakers intuitively and holistically. As a result, there was a significant difference between the pre- and post-tests in the reading test and the listening test, although there was no significant difference in the interview. The results showed that the chanting method helped to improve the pronunciation and listening abilities of the participants at the time.
However, this improvement in their pronunciation ability was temporary and had not become an automatic feature of their skills.

Also, Kawai (2014) conducted a study using jazz chants with 52 Japanese college students. The participants were classified into three different English-proficiency groups based on their scores in the TOEIC Bridge test, $F (2, 49) =59.8$, $p < .001$, and $\omega^2 = .69$. Over a five-week period, participants were asked to chant the story while listening to the CD for about 15 minutes at the beginning of every weekly class. 12 test sentences (ISI1–ISI4) quoted directly from Nakano (2007) were randomly arranged. This study used “changes” in the ISI duration between the pre- and post-tests as the criteria for acquiring English stress-timed rhythm in speech production. The idea was based on earlier research results from Mochizuki-Sudo and Kiritani (1991), Nakano (1997, 2007), and Kohno (2001). The results showed that the ISI duration of all three groups reduced significantly. In particular, there was a big difference between pre- and post-test data of the low English-proficiency group. It concluded that the practice of jazz chants was effective, especially for students in the low English-proficiency group in improving their pronunciation, listening ability, and motivation to learn English.

2.3 The Present Study

According to the results of the previous research above, jazz chant instruction significantly helped to improve the participants’ ability in pronunciation and speaking. However, these were immediate effects, and long-term research is needed. Moreover, it is difficult to make direct comparisons because the number of participants and the test methods were different in each research project. In the study by Kawai (2014), the ISI duration between pre- and post-test decreased significantly; however, the author believes it is more important for raters to realize the progress or difference between pre- and post-acoustic data when they actually listen to them. This thinking has led to the present study.

The research questions are as follows:

**RQ1:** Are there any differences between pre- and post-test scores as evaluated by the human ear?

**RQ2:** Does the change in ISI duration correlate with the scores evaluated by the human ear?

The answers to RQ1 and RQ2 could be related to two focus points: How does a chant method work as an instruction in pronunciation for college students? What elements of Japanese EFL learners’ English pronunciation influence native English speakers’ ratings after receiving instruction in chanting?

3. Methods

3.1 Data collection

In this present study, the author extracted 160 parts of acoustic data from the low English-proficiency group in Kawai (2014). See Table 1. In Kawai (2014), she analyzed and
reported only the mean differences of four sentence types (ISI1–ISI4) between the pre- and post-tests for each set. In this study, the author analyzes and reports on the results of these four test sentences from one set in detail.

3.2 Raters

The raters were two native English speakers. One is a Canadian (male, 68 years old) and the other is an American (female, 44 years old). Both of them are experienced EFL teachers and have been teaching in various kinds of institutions such as kindergarten, high schools, and universities in Japan for many years. The inter-rater reliability of two raters was confirmed by Cronbach’s coefficient alpha (=.88). It proved to be sufficiently reliable.

3.3 Procedure and Analysis

These 160 samples (pre- and post-test) above were mixed on one CD. Before the formal ratings, the raters and the author had a meeting and some training time. The author asked them to focus on prosodic factors, especially how natural the samples are in terms of English stress-timed rhythm, and to evaluate them holistically and intuitionally by using a 9-point scale (1: very unnatural, 9: very natural). Then, we evaluated three samples (included in the data) as a practice and discussed them individually. Later, they evaluated the data formally. According to Jesney (2004), many researchers use global accent settings in such a way that native speakers judge non-native speakers’ speech holistically and intuitionally by using a 3–10-point Likert scale.

In the analysis, two-way repeated measures ANOVA was conducted with pre- and post-test results. The within-subjects-factors were sentence type S1–S4 (ISI1–ISI4) and time (pre-and post-test). Partial eta-squared ($\eta^2$) was used as effect sizes. In addition, in order to investigate the correlation between the ISI duration and two raters’ scores, one-way ANOVA was conducted and correlation ratios were used as index instead of correlation coefficient because there is qualitative difference between ISI duration and two raters scores.

Table 1

<table>
<thead>
<tr>
<th>Participants</th>
<th>High level</th>
<th>Middle level</th>
<th>Low level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic data</td>
<td>(2014) 12 sentences $\times$ 52 $\times$ 2 (pre-and post-test) = 1,248 sentences</td>
<td>(this study) 4 sentences $\times$ 20 $\times$ 2 (pre-and post-test) = 160 sentences</td>
<td>20 (in Author, 2014)</td>
</tr>
<tr>
<td>Materials</td>
<td>Jazz Chant Fairy Tales (Graham et al., 1988) + CD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period</td>
<td>15 minutes $\times$ 5 times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>ISI duration (MS) measured by Sugi Speech Analyzer</td>
<td>(this study) 2 English native teachers (a Canadian, an American)</td>
<td></td>
</tr>
</tbody>
</table>
4. Result

4.1 The result of the acoustic analysis

Tables 2 shows the descriptive statistics on the pre- and post-test. Table 3 shows the mean differences between the pre- and post-test. As shown in Table 4 and Figure 1, there was a significant interaction between sentence type and time ($F (1.9, 36.5) = 10.62, p < .001$, \(\text{partial } \eta^2 = .36\)). The main effect for sentence type ($F (3, 57) = 629.31, p < .001$, \(\text{partial } \eta^2 = .97\)) and the main effect for time ($F (1,19) = 157.87, p < .001$, \(\text{partial } \eta^2 = .89\)) were also observed. And, the simple main effect for sentence was significant in all sentence types (\(S_1 F (1, 19) = 17.93, p < .001$, \(\text{partial } \eta^2 = .49\); \(S_2 F (1, 19) = 31.85, p < .001$, \(\text{partial } \eta^2 = .62\); \(S_3 F (1, 19) = 45.92, p < .001$, \(\text{partial } \eta^2 = .71\); \(S_4 F (1, 19) = 75.94, p < .001$, \(\text{partial } \eta^2 = .80\).)

Table 2

<table>
<thead>
<tr>
<th>Sentence</th>
<th>N</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>S1</td>
<td>20</td>
<td>672.80</td>
<td>120.26</td>
</tr>
<tr>
<td>S2</td>
<td>20</td>
<td>1004.45</td>
<td>120.19</td>
</tr>
<tr>
<td>S3</td>
<td>20</td>
<td>1272.15</td>
<td>121.28</td>
</tr>
<tr>
<td>S4</td>
<td>20</td>
<td>1675.35</td>
<td>170.46</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Sentence</th>
<th>N</th>
<th>Pre-Post</th>
<th>SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>20</td>
<td>108.80</td>
<td>114.92</td>
<td>[55.02, 162.58]</td>
</tr>
<tr>
<td>S2</td>
<td>20</td>
<td>161.65</td>
<td>128.10</td>
<td>[101.70, 221.60]</td>
</tr>
<tr>
<td>S3</td>
<td>20</td>
<td>224.90</td>
<td>148.42</td>
<td>[155.44, 294.36]</td>
</tr>
<tr>
<td>S4</td>
<td>20</td>
<td>350.70</td>
<td>179.98</td>
<td>[266.47, 434.93]</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; LL = lower limit, UL = upper limit.

Table 4

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>(\text{partial } \eta^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence type</td>
<td>3</td>
<td>565718.69</td>
<td>730.29</td>
<td>&lt;.001</td>
<td>.98</td>
</tr>
<tr>
<td>Time</td>
<td>1</td>
<td>1834623.06</td>
<td>184.37</td>
<td>&lt;.001</td>
<td>.91</td>
</tr>
<tr>
<td>Sentence × Time</td>
<td>1.93</td>
<td>149166.48</td>
<td>9.54</td>
<td>&lt;.001</td>
<td>.33</td>
</tr>
</tbody>
</table>
However, it is not clear how much ISI durations were reduced in a whole sentence between pre- and post-test. To investigate the change of duration of the ISI in whole sentences, Table 6 shows the ratio of the ISI duration (%) in whole sentences’ duration. Two-way repeated measures ANOVA was conducted with pre- and post-test results. The within-subjects-factors were sentence type S1–S4 (ISI1–ISI4) and time (pre- and post-test). As shown in Table 7, there was a significant main effect for sentence type \( (F(3,57) = 236.78, p < .001, \text{partial } \eta^2 = .93) \). However, the main effect for time \( (F(1,19) = 3.00, \text{ns}) \) and the interaction between sentence type and time \( (F(3,57) = 2.64, \text{ns}) \) were not observed. After the multi comparison (Bonferroni), the differences among ISI ratio of all sentences were significant at 5% level.

### Table 5

**Mean Duration between Pre-test and Post-test of Whole Sentence (MS)**

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Pre-test ((SD))</th>
<th>Post-test ((SD))</th>
<th>Pre–Post ((SD))</th>
<th>95% CI (Pre-Post)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1530.85 ((174.95))</td>
<td>1284.35 ((170.82))</td>
<td>246.50 ((201.96))</td>
<td>[151.98, 341.02]</td>
</tr>
<tr>
<td>S2</td>
<td>1915.00 ((152.50))</td>
<td>1572.45 ((200.20))</td>
<td>342.55 ((199.31))</td>
<td>[249.27, 435.83]</td>
</tr>
<tr>
<td>S3</td>
<td>2062.65 ((211.78))</td>
<td>1764.10 ((214.37))</td>
<td>298.55 ((245.31))</td>
<td>[183.74, 413.36]</td>
</tr>
<tr>
<td>S4</td>
<td>2558.40 ((284.95))</td>
<td>2093.30 ((276.35))</td>
<td>465.10 ((291.24))</td>
<td>[328.80, 601.40]</td>
</tr>
</tbody>
</table>

*Note.* CI = confidence interval; LL = lower limit, UL = upper limit.

### Table 6

**Mean Duration of ISIs /Mean Duration of a Whole sentence (%)**

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Pre-test ((SD))</th>
<th>Post-test ((SD))</th>
<th>Pre–Post ((SD))</th>
<th>95% CI (Pre-Post)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>43.68 ((4.67))</td>
<td>43.22 ((4.77))</td>
<td>.46 ((5.35))</td>
<td>[–2.04, 2.96]</td>
</tr>
<tr>
<td>S2</td>
<td>52.39 ((3.58))</td>
<td>53.24 ((4.18))</td>
<td>–.85 ((3.28))</td>
<td>[–2.38, .69]</td>
</tr>
<tr>
<td>S3</td>
<td>61.84 ((4.14))</td>
<td>59.50 ((5.08))</td>
<td>2.34 ((4.31))</td>
<td>[.33, 4.36]</td>
</tr>
<tr>
<td>S4</td>
<td>65.72 ((5.00))</td>
<td>63.54 ((3.85))</td>
<td>2.18 ((4.74))</td>
<td>[–.05, 4.40]</td>
</tr>
</tbody>
</table>

*Note.* CI = confidence interval; LL = lower limit, UL = upper limit.

As shown in Table 5, the participants compressed whole sentences duration much more in S4 (ISI4) than in S1(ISI1). However, according to Table 6, as the number of ISIs increased, each

![Figure 1. Change of ISI duration in pre-and post-test.](image-url)
ratio of the ISI duration (%) in whole sentences increased in pre-and post-test. The result shows that as the number of ISIs increased as it became more difficult for the participants to compress ISI duration in reading the test sentences. Also, ISI ratios in each whole sentence were not so different between pre-and post-test because the main effect for time was not observed. Consequently, the result shows that the participants compressed not only ISI duration but also whole sentences duration in post-test than in pre-test in the study.

Table 7
Tests of Within-Subjects Effects (ISI)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence type</td>
<td>3</td>
<td>3498.55</td>
<td>236.78</td>
<td>&lt; .001</td>
<td>.93</td>
</tr>
<tr>
<td>Time</td>
<td>1</td>
<td>42.46</td>
<td>3.00</td>
<td>.100</td>
<td>.14</td>
</tr>
<tr>
<td>Sentence × Time</td>
<td>3</td>
<td>22.88</td>
<td>2.64</td>
<td>.058</td>
<td>.12</td>
</tr>
</tbody>
</table>

4.2 The result of evaluation by the human ear

Table 8 and Table 9 show the descriptive statistics of two raters’ scores in the pre- and post-test. A full score is 18 points ($9 \times 2 = 18$). Table 10 shows the mean differences between the pre- and post-test. As shown in Table 10, the main effect for sentence type ($F (2.1, 39.5) = 20.91$, $p < .001$, partial $\eta^2 = .52$) and the main effect for time ($F (1, 19) = 26.51$, $p < .001$, partial $\eta^2 = .58$) were observed. However, there was not a significant interaction between sentence type and time ($F (1.8, 33.5) = .27$, ns.). After the multi comparison (Bonferroni), the differences of evaluation scores among S1–S2, S1–S3, S1–S4 were significant at 1% level, while that of S2–S3, S2–S4, S3–S4 were not significant. Figure 2 shows that the scores decreased as the number of ISI increased. The significant differences among S1–S2, S1–S3, and S1–S4 by multi comparison (Bonferroni) also show the same tendency. Similar to the result of ISI duration, the result of the evaluation by human ears shows that it became more difficult for the participants to read the test sentences as the number of ISI increased. However, the raters scores were higher in post-test than in pre-test significantly. Accordingly, these result above shows that the participants’ speech rate changed significantly after the chanting practice. They read the test sentences much faster than before.

Figure 2. Change of two raters’ scores in pre-and post-test.
Table 8

Mean Scores by Two Raters in the Pre-and Post-test (9-point scale; full score = 18)

<table>
<thead>
<tr>
<th>Sentence</th>
<th>N</th>
<th>Pre-test Mean</th>
<th>SD</th>
<th>Post-test Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>20</td>
<td>6.00</td>
<td>3.01</td>
<td>7.80</td>
<td>2.33</td>
</tr>
<tr>
<td>S2</td>
<td>20</td>
<td>4.05</td>
<td>1.50</td>
<td>6.00</td>
<td>2.75</td>
</tr>
<tr>
<td>S3</td>
<td>20</td>
<td>4.35</td>
<td>1.50</td>
<td>5.60</td>
<td>2.06</td>
</tr>
<tr>
<td>S4</td>
<td>20</td>
<td>3.65</td>
<td>.99</td>
<td>5.25</td>
<td>1.94</td>
</tr>
</tbody>
</table>

Table 9

Mean Scores by Two Raters between Pre- and Post-test (9-point scale; full score = 18)

<table>
<thead>
<tr>
<th>Sentence</th>
<th>N</th>
<th>Pre-Post</th>
<th>SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>20</td>
<td>−1.80</td>
<td>3.05</td>
<td>[−3.23, −.37]</td>
</tr>
<tr>
<td>S2</td>
<td>20</td>
<td>−1.95</td>
<td>3.09</td>
<td>[−3.39, −.51]</td>
</tr>
<tr>
<td>S3</td>
<td>20</td>
<td>−1.25</td>
<td>2.49</td>
<td>[−2.42, −.08]</td>
</tr>
<tr>
<td>S4</td>
<td>20</td>
<td>−1.60</td>
<td>2.28</td>
<td>[−2.67, −.53]</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; LL = lower limit, UL = upper limit.

Table 10

Tests of Within-Subjects Effects (two raters scores)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence type</td>
<td>3</td>
<td>66.18</td>
<td>20.91</td>
<td>&lt; .001</td>
<td>.52</td>
</tr>
<tr>
<td>Time</td>
<td>1</td>
<td>107.26</td>
<td>26.52</td>
<td>&lt; .001</td>
<td>.58</td>
</tr>
<tr>
<td>Sentence $\times$ Time</td>
<td>1.76</td>
<td>1.40</td>
<td>.27</td>
<td>.741</td>
<td>.01</td>
</tr>
</tbody>
</table>

4.3 The correlation between acoustic analysis and evaluation by the human ear

The correlation ratios in pre-test were $F(1, 38) = 614.44, p < .001, \eta = .97$ in S1, $F(1, 38) = 1385.18, p < .001, \eta = .99$ in S2, $F(1, 38) = 2185.21, p < .001, \eta = .99$ in S3, and $F(1, 38) = 1923.49, p < .001, \eta = .99$ in S4. Respectively, those in post-test were $F(1, 38) = 708.35, p < .001, \eta = .97$ in S1, $F(1, 38) = 937.60, p < .001, \eta = .98$ in S2, $F(1, 38) = 822.87, p < .001, \eta = .98$ in S3, and $F(1, 38) = 1094.06, p < .001, \eta = .99$ in S4. It is clear that the correlation ratios in each sentence type between the ISI duration and two raters scores were high. However, it is unclear whether there is a negative correlative tendency between two factors. Hence, the author shows the scatter plots of each sentence set in pre-and post-test. See Figure 3. According to the analysis and Figure 3, the tendency of negative correlation between the result of the ISI duration and two raters scores was observed.
4.4 The answers to the RQs

RQ1: Are there any differences between pre- and post-test scores as evaluated by the human ear?

The answer is “yes.” The two raters gave higher scores in the post-test than in the pre-test when they evaluated these acoustic data, and their inter-rater reliability was highly significant. As a result, significant differences were observed between the pre- and post-test in all sentence types S1 (ISI1)–S4 (ISI4), respectively, through two-way repeated measures ANOVA in the study.

RQ2: Does the change in ISI duration correlate with scores evaluated by the human ear?

The answer is “yes.” In the study, the correlation ratios between ISI duration and the scores by human ear were high. Two raters score increased as the ISI’s duration decreased. The negative correlation tendency between the ISI duration of acoustic analysis and the scores by the human ear evaluation was observed through the scatter plots and the correlation ratios.

5. Discussion

5.1 Chants as pronunciation practice

Szpyra-Kozlowska (2014) mentioned that pronunciation is regarded as the most difficult aspect to master when learning another language. However, in this study, significant differences were observed between the pre- and post-test: the scores of two raters were significantly higher in the post-test than in the pre-test. This result supports the previous research (Nakano, 1997; Kawai, 2009; 2014; Kung, 2013). This means that intensive chanting practice with rhythm promoted more fluency on their pronunciation in adult EFL learners, just as Celce-Murcia et al. (1996)
pointed out that jazz chants is a good way to train learners to read English texts more fluently rather than word by word. The result of the study supports their idea, and chanting practice with rhythm can be a way to develop Japanese adult EFL learners’ pronunciation.

In terms of improving the intelligibility of pronunciation, the issue of segmental vs. suprasegmental has been under discussion for a long time. Pennington (1987) mentioned the priority of the suprasegmental aspect and suggests the usefulness of a top-down approach in both pronunciation instruction and listening instruction. Morley (1987) suggested that teachers should teach the suprasegmental aspect first and the segmental aspect next, such as vowels and consonants. Such a way of thinking seems to be more natural, considering the order of acquisition in the mother tongue.

However, Szpya-Kozlowska (2014) stated that neither of these two perspectives can be considered fully adequate, as phonetic priorities cannot be established for all learners of different linguistic backgrounds but must be specified for each L1 on the basis of the phonetic distance between the L1 and L2. This should be coupled with empirical research into the role of individual phonetic features as regards safeguarding intelligibility (p.230). In addition, Szpya-Kozlowska (2013) argued that pedagogical priorities can only be established when different variables of the teaching process are taken into consideration, such as its participants and their linguistic background, their specific needs and problems, quoting Celce-Murcia et al (1996).

Focusing on native Japanese speakers, Sugito (1996) and Sudo (2010) explained a characteristic of Japanese EFL learners’ English pronunciation: most Japanese EFL learners tend to stress every individual word when they read English texts. Likewise, the raters in the study, two native English speakers, pointed out over-stress and “katakana English” as a characteristic of Japanese EFL learners. That tendency is attributed to the big difference between Japanese mora-timed rhythm and English stress-timed rhythm and acquiring an English stress-timed rhythm is acknowledged as one of the most difficult aspects for Japanese EFL learners. Kenworthy (1987) analyzed the issues of Japanese EFL learners’ pronunciation and suggested the importance of teaching the suprasegmental aspect, especially rhythm. According to 2.1, the present situation in teaching pronunciation in Japanese EFL classrooms has not improved a great deal compared to 30 years ago. Moreover, these problems are frequently pointed out in both EFL and ESL settings across the world as mentioned: pronunciation instruction is neglected or less regarded; there is no clear goal in teaching pronunciation, and teachers lack confidence in teaching pronunciation.

However, allocating plenty of classroom time for pronunciation practice is unrealistic given the importance of grammar and vocabulary in improving students’ speaking ability. The author suggests “shadowing with jazz chants or other chants” as one classroom activity to promote pronunciation practice. It is teacher-friendly and very easy to use in class. Even in a short time, it can make immediate improvements to Japanese adult EFL learners’ pronunciation. If continued sustainably, it can provide good training for Japanese adult EFL learners to improve their fluency.
by accelerating their speech rate and improving other aspects of prosody, such as rhythm, or their pronunciation. In addition, teachers can use it at the beginning of class or at any time they want to change the atmosphere in the classroom because the music and rhythm of chants can make the classroom happier and more pleasant. As Richard-Amato (2003) mentioned, chanting English texts aloud with the beat of a drum helps students reduce their anxiety in English classrooms. Guiora (1972) suggested that there is a close relationship between learners’ pronunciation and affective factors because pronunciation is the most critical factor in a learner’s identity; furthermore, improving pronunciation is important for increasing confidence and successfully communicating in English. Recently, Renner et al. (2017), following Guiora (1972), indicated that acute alcohol consumption may have beneficial effects on the pronunciation of a foreign language in people who have recently learned that language.

5.2 Correlation between ISI duration and human ear evaluation

In this study, the negative correlation between ISI duration and human ear evaluation was observed. Immediately after an intensive practice using jazz chants, most of the participants could read the test sentences much faster than before. The phenomenon was also observed in an acoustic data that demonstrated decreased ISI durations. In the same way, two raters of the study felt that the output was more natural when they listened to and evaluated the same acoustic data in the post-test than in the pre-test. When ISI’s duration decreased, the two raters’ scores become higher. Participants’ improved fluency, such as speech rate and rhythm, affected the raters’ evaluation scores. Appropriate speech rate and keeping a good rhythm can be said to be factors that are attributed to the intelligibility of pronunciation.

Mochizuki et al. (1991) reported that when unstressed syllables are added (ISI1→ISI4), low English-proficiency learners did not control ISI duration well, although high English-proficiency learners did. In Nakano’s (1997) study, the same tendency was observed in the pre-test. However, after rhythm instruction using jazz chants, the low English-proficiency learners in her research compressed the ISI duration almost the same as the high English-proficiency learners. This study also followed the result and the ISI duration decrease between pre-and post-test. The effect of reading practice using jazz chants was observed. However, in this study, the raters scores became lower generally as unstressed syllables were added (ISI1→ISI4) despite ISI duration much decreased between pre-and post-test. Nakano (1997) reported that when unstressed syllables were added (ISI1→ISI4), it was difficult for all the participants to compress the weak vowels of ISI compared to those of native English speakers. Previous studies and the result of this study show that it seems difficult for Japanese native speakers to fluently read or produce English sentences or phrases that include more than two unstressed syllables. In the result of this study, each effect size in the simple main effect for sentence became larger when unstressed syllables increased (ISI1→ISI4). Reading by using jazz chants might be one of good practice for reducing ISI durations.
6. Conclusion and limitation of this study

In the study, the participants’ pronunciation significantly improved in the post-test after chanting practice with jazz chants. Moreover, the tendency that correlation between ISI duration and the scores judged by two native English speakers was confirmed in the study. The results indicate that 1) chanting practice with jazz chants can be a good approach for Japanese adults EFL learners to improve their pronunciation, 2) change of speech rate and ISI duration affect native English speakers’ judgment, and 3) appropriate speech rate and a good rhythm are attributing factors to the intelligibility of pronunciation.

Although its positive effects were confirmed immediately after the chanting practice, long-term experiments and revalidation after a long interval, such as six months, are needed. In addition, one of the raters mentioned that lack of context in the material made the evaluation challenging and that some of the utterances were unnatural in spoken English, and therefore, the stress pattern was not readily apparent. Hence, the test material also needs to be reconsidered.

Learning and acquiring correct pronunciation is a very complex experience, and it takes time. However, clear and intelligible pronunciation helps EFL learners to communicate smoothly with each other and with native speakers, and it promotes their confidence. It is a very important role for English teachers to keep motivating learners and fostering learners’ confidence.

Notes
1. The idea was referred to by Guiora (1972). It is almost impossible to change pronunciation once ego development is complete, but the goal can still be achieved by fostering self-permeability.

Acknowledgement

I would like to give my great thanks for the participants and raters. Also, I really thank the anonymous reviewers for their helpful comments. This work was supported by JSPS KAKENHI Grant Numbers JP15K02689 and JP18K00872.

References


doi: 10.1111/j.1467-1770.1995.tb00963.x


Teshima, R. (2011). Nihon no chugakko kotogakko ni okeru Eigo no onseikyoiku ni tsuite: hatsuo shido no genjo to kadai [English pronunciation instruction in junior and senior high schools in Japan: The present situation and issues]. *Onseikenkyu. 15*, 31–43. doi: 10.24467/onseikenkyu.15.1_31