Effects of Change in Speech Rates on Japanese EFL Listeners’ Word Recognition

Hirokazu YONEZAKI
Kindai University

Abstract

The main focus of the present study is to investigate whether high-speed listening in English class is effective in enhancing word recognition of elementary-level Japanese EFL listeners. Both experimental and anecdotal evidences suggest that, with continuous exposure to a faster speech rate, listeners can normalize a faster speech rate as their new baseline and find the original speech rate slower. In order to confirm if steady exposure to a faster speech rate has a positive effect on the listener’s word recognition as well as comprehension, an experiment was conducted. In the experiment, participants in the three experimental groups were exposed to compressed speech rates at 1.2, 1.5, and 2.0 times the original rate for half a year in class, while those in the control group listened at the original rate. The results of the posttests showed that the participants who listened at 1.5 times the original rate fared significantly better than the other three groups in terms of word recognition. However, no difference between the groups was found in terms of comprehension.

1. Introduction

The aim of the present study is to investigate the effects of training elementary-level Japanese EFL listeners with different speech rates on their word recognition. English is a stress-timed language, the prosody of which is very different from that of Japanese, a syllable-timed one. In a stress-timed language, articulation of function words are shorter and quicker with occasional omission than that of content words (Ur, 1984; Rost, 2002). Yonezaki (2016) showed that the recognition of function words was far more difficult for elementary to intermediate level listeners to recognize than that of content words and that speech rates have an important role to play in the recognition gap between the two word categories, depending on the listeners’ proficiency levels. In addition, both anecdotal and experimental studies showed that change of speech rates have significant effects on L2 learners’ listening comprehension (Dupoux & Green, 1997). In light of this background, an experiment was conducted to see how long-term listening sessions on the same materials with different speech rates affects L2 learners’ listening proficiency, especially the first phase of listening: lexical identification from the auditory input.
1.1 Background

The focus of English education in Japan has long been placed on improving reading skills rather than listening. In the EFL environment, therefore, it is quite natural that they have had fewer opportunities where they listen to spoken language than read written language (Yamaguchi, 1997). Moreover, Japanese EFL learners are said to rely more on visual language than auditory information when they try to comprehend the incoming linguistic message (Ito, 1989; Yamaguchi, 1999; Yanagawa, 2016). Their dependence on visual rather than auditory stimuli, coupled with infrequent learning opportunities of basic vocabulary through listening (Yamaguchi, 1997), makes it challenging for Japanese EFL learners to recognize spoken words without the help of written script, which is one of the main reasons why quite a few Japanese EFL learners do not feel confident about listening.

1.2 Listening Process and Japanese EFL Learners’ Weak Point

According to Vandergrift and Goh (2012), listening process involves three main phases: perception, parsing, and utilization. Perception involves the recognition of incoming sound signals by the listener as words or meaningful chunks of the language. The perception phase involves bottom-up processing and depends on such factors as speed of the sound stream. On the other hand, parsing involves the segmentation of an utterance according to syntactic and semantic cues, creating mental representation of the combined meaning of the words. Both bottom-up and top-down processing are involved while the parser attempts to segment the sound stream into meaningful units, through phonological analysis and word retrieval from the mental lexicon. Finally, utilization, which is top-down in nature, involves creating mental representation of what is retained by the perception and parsing processes and linking this to existing knowledge stored in long-term memory.

Of these phases, what Japanese EFL learners find most challenging is the phases from perception to parsing, especially the segmentation of an utterance, informed of by perception, into meaningful chunks of words, that is, spoken word recognition (Ito, 1989; Hayashi, 1991; Yamaguchi 1997). Ito (1989) holds that their auditory vocabulary is much smaller than their visual vocabulary. If Japanese EFL learners cannot comprehend an utterance without the written script, their grammatical and lexical knowledge are of little use in listening. Hayashi (1991) claims that failure to comprehend even at a sentential level may be caused by inefficient processing of individual words. Noro (2006) also states that unfamiliarity with native speakers’ pronunciation is yet another big hurdle, which leads to Japanese EFL learners’ failure to recognize words.

There is no question, therefore, about the claims that one of the important roles of listening instruction is helping learners deconstruct speech in order to recognize words and phrases quickly (Vandergrift, 2007), and that the problem in listening is how to match unintelligible chunks of language with their written forms (Field, 2008). In addition, when bottom-up processing is accurate and automatic, it frees working memory capacity and thus allows the listener to build
complex meaning representations. However, when it is not, it may limit the listener’s ability to form a detailed and coherent message (Field, 2008).

In locating lexical boundaries and identifying words, one of the most challenging factors for L2 learners of English is that L1 speakers’ natural speech is simply too fast (Griffiths, 1990; Graham, 2006; Noro, 2006). Speech rate is closely related to bottom-up processing in the perception phase (Vandergrift & Goh, 2012), and the spoken text delivered at natural speed makes lexical boundaries in continuous sound streams blurry and ambiguous, making word recognition, especially that of function words, more challenging for L2 listeners (Goh, 2000).

Especially, Japanese EFL learners have been accustomed to listening to the slower rate of speech, which makes speech rate the greatest source of their listening problems (Noro, 2006). If speech rate is too fast or natural for L2 listeners, the first phase of listening, perception, followed by lexical segmentation and word recognition, will not work (Yanagawa, 2016).

1.3 Speech Rate Manipulation and Listening Comprehension

According to Adank and Devlin (2010), when L1 speakers listened to compressed speech up to 45%, about 1.8 times the original rate, their adaptation-related changes were observed in four separate areas both in left and right hemisphere of the brain. The changes in the right hemisphere means that adaptation may have occurred at an acoustic, rather than linguistic, level. In contrast, the changes in the left hemisphere means that adaptation-related changes have occurred at a linguistic level; that is, these changes are related to comprehending the speech. They showed that, after hearing just 16 sentences, L1 listeners’ comprehension became faster and more accurate to time-compressed speech up to 1.8 times the original rate.

Poldrack et al. (2001) also suggested that activation in the left superior temporal sulcus (STS), which is related to linguistic cues, increased as sentence compression rate rose up to 3.3 times the original rate.

According to Dupoux and Green (1997), L1 listeners adapted to sentences about 2.6 times the original rate. They claim that the perceptual system alters its criteria for judging such incoming cues in relation to the rate at which the speech was produced. They argue that some of the improvement in the ability to recognize compressed speech is due to perceptual mechanisms involved in the normalization for speech rate. In other words, some kind of low-level tuning of the perceptual system, or perceptual recalibration or adjustment, to accommodate faster speech rate has occurred (Dupoux & Green, 1997; Golomb, Peelle, & Wingfield, 2007).

Kajiura (2016) showed that Japanese EFL learners made improvements in their listening proficiency after five-day practice, which sums up to about fourteen hours, with faster rate of speech. She gives two reasons why L2 learners might be able to become accustomed to listening at a fast speech rate with steady practice. One is that practice with faster speech makes it possible to understand the faster speech itself. The other is that listening practice at a faster speech rate makes learners capable of processing sounds and meanings more quickly.
As to whether the adjustment to fast rate of speech is maintained for a long term despite the need to adjustment to other slower rates, which means that the adjustment is related to permanent learning, or whether the normalization resets after the presentation of uncompressed materials, Dupoux & Green (1997) argue that presentation of slower rates does not cause a complete resetting of adjustment parameters to baseline, which means that some kind of perceptual learning takes place. This argument is supported by Golomb et al. (2007), who claim that adaptation to time-compressed speech is not hindered by brief pauses or the insertion of uncompressed sentences.

These studies suggest that listeners are able to make perceptual adjustment to highly compressed speech, both acoustically and linguistically, after several sessions of training, that the effect of adaptation is not disrupted by the normal rate of speech, and that the effect lasts for a comparatively longer period of time.

**1.4 Focus of the Present Study**

As stated above, there is enough experimental evidence that listeners can normalize a faster speech rate as their new baseline. In addition, according to Dupoux and Green (1997), anecdotal evidence is also abundant that, with continuous exposure to a faster speech rate, listeners find it easier to understand it and report that it sounds less fast. That is to say, when listeners make perceptual adjustment to a faster rate of speech, some kind of normalization of a faster rate takes place, which renders the original baseline somewhat slower rate to the listener.

However, empirical evidences concerning the effects of speech rate manipulation on L2 listeners is scarce. If L2 listeners, especially listeners of lower proficiency, can adapt to faster speech rate in perceptual phase, it may affect their listening positively, because, after steady practice with faster speech, they now feel that the original rate is slow. This can best be expected in the first phase of listening, perception, which may lead to improvement in word recognition because improvement in perceptual phase may have some positive effect on the second phase, parsing. Through the improvement in bottom-up processing, some positive effect on comprehension may also be expected.

Therefore, in the present study, the following three research questions (RQs) are addressed:

RQ1: Is steady exposure to a faster rate of speech in class effective in improving word recognition by Japanese EFL listeners with lower levels of proficiency?
RQ2: Is there any specific speech rate at which low-proficiency listeners fare best in word recognition after the exposure to the faster speech rate?
RQ3: Does steady exposure to a faster speech rate have any positive effect on listening comprehension as well?
2. Method

2.1 Participants
The participants were 206 first-year technical college students in Japan who majored in engineering. Their L1 is Japanese. The participants ranged in their levels of English proficiency from a beginner to an elementary level; their mean score of the TOEIC bridge test was 108.88 (SD = 10.72).

2.2 Materials
The experiment had a pretest-treatment-posttest design. For the pretest and the posttest, both multiple-choice listening comprehension tests and word recognition tests, which had a cloze-test format, were conducted. For the listening comprehension tests, pre-second grade STEP (Society for Testing English Proficiency, 2009 & 2010) listening tests consisting of 30 questions each were adopted. For the word recognition tests as well, pre-second grade STEP listening tests (2008 & 2009), which were different from the ones used for the listening comprehension tests, were adopted. Materials for the posttests were different from the ones for the pretests. However, they shared the same format.

The word recognition tests were transcription tests and the participants were required to write down one word in each blank, which they thought they had heard. For both the pretest and the posttest, five sections each for the dialogue and the monologue part of the STEP test, ten in total, were extracted. In these ten sections, 25 words each in the dialogue and the monologue, 50 in total, were blanked out. Of the 50 words which were blanked out, 25 were content words with the rest being function words. Quirk, Greenbaum, Leech, and Svartvik (1985) were referred to in distinguishing the two word categories.

All the 50 blanked-out words in both the pretest and the posttest were at a junior-high school level in Japan. The participants would have had little difficulty in recognizing them, if those words had been in written scripts or articulated individually.

As to the material used in the treatment, one of the English textbooks authorized by MEXT, Vision Quest vol. 1, published by Keirinkan was adopted, because this was the only and main textbook that the participants used in class.

2.3 Procedure
First, 206 participants were divided into four groups: one control group (Control Group) and three experimental groups (Experimental Groups 1, 2, and 3), who listened to the treatment materials at different speech rates. Participants in Control Group listened at the original speech rate, while those in the experimental groups at mechanically compressed speech rates. Experimental Group 1 listened at 1.2 times the original rate, Experimental Group 2 at 1.5 times, and Experimental Group 3 at 2.0 times. The 206 participants were made up of five classes and the
groups were divided according to the classes they belonged to. Consequently, two of the five classes belonged to Experimental Group 2, while the other groups were each made up of one class.

In compressing the speech rates, a speed-changing software attached by the publisher was used. Table 1 shows the respective number of participants in each group and the average speech rates of the pretest, posttest, and treatment materials, listened by four groups, in words per minute (wpm). The participants listened to all the dialogues, example sentences in the grammar sections, and exercises in the textbook at the manipulated speed for each group.

Table 1

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>Listening Comprehension Tests</th>
<th>Word Recognition Tests</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pre</td>
<td>Post</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>42</td>
<td></td>
<td></td>
<td>126</td>
</tr>
<tr>
<td>Experimental 1</td>
<td>41</td>
<td>129</td>
<td>113</td>
<td>151</td>
</tr>
<tr>
<td>Experimental 2</td>
<td>83</td>
<td></td>
<td>131</td>
<td>189</td>
</tr>
<tr>
<td>Experimental 3</td>
<td>40</td>
<td></td>
<td></td>
<td>252</td>
</tr>
</tbody>
</table>

The treatment lasted about half a year until the posttest was conducted. During the treatment period, the participants were given normal classes, one 90-minute session a week, in which they learned English, using the textbook. They had 15 sessions during the half-a-year period. Accordingly, the only difference between the groups was the speeds at which they listened to the CD attached to the textbook. The activities using the CD at the manipulated speed involved listening practice with the textbook closed, listening with the textbook open and with the participants following the written scripts while listening, a couple of sets of sentence-by-sentence repeating and shadowing practice, reading the Japanese translation while listening, and shadowing practices with the participants looking at the Japanese translation. All these activities involved were given both before and after explicit syntactical and phrasal explanation concerning the scripts.

As far as word recognition tests are concerned, each of the 50 (25 content and 25 function) words in the pretest and the posttest were graded either correct or incorrect. The participants were discouraged from using *katakana* when they were unsure of the spellings, and asked to use alphabet letters that they thought they had heard. Therefore, all the *katakana* answers were judged to be incorrect.

In grading, if the sound was recognized correctly, the item was judged to be correct, even if it was misspelled. However, if, for example, *here* was misspelled as *hear* or *buy* as *by*, they were judged to be incorrect. These misspellings may well have been caused by a failure to recognize a word at the parsing phase and accurate word recognition must involve both perception and parsing phases while the listener segments the incoming speech, through phonological analysis and word
retrieval from the listener’s mental lexicon, into meaningful units, with reference to syntactic and semantic cues (Vandergrift & Goh, 2012).

All the data were computed into the percentage of accurate word recognition, with the number of items correctly recognized being the numerator and the total number of blanked-out items the denominator. Content words and function words were separately examined.

3. Results

3.1 Word Recognition Tests

Table 2 shows the descriptive statistics of the words correctly recognized in percentage in the pretest (Cronbach’s alpha = .789) and in the posttest (Cronbach’s alpha = .828).

Table 2
Descriptive Statistics of the Pretest and the Posttest (Words Correctly Recognized, %)

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Pretest Mean (SD)</th>
<th>Posttest Mean (SD)</th>
<th>Total Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Content</td>
<td>Function</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>42</td>
<td>39.43 (13.31)</td>
<td>39.05 (11.51)</td>
<td>39.24 (11.14)</td>
</tr>
<tr>
<td>Experimental 1</td>
<td>41</td>
<td>34.83 (14.63)</td>
<td>37.76 (13.11)</td>
<td>36.29 (12.85)</td>
</tr>
<tr>
<td>Experimental 2</td>
<td>83</td>
<td>39.98 (13.79)</td>
<td>37.83 (14.43)</td>
<td>38.46 (13.05)</td>
</tr>
<tr>
<td>Experimental 3</td>
<td>40</td>
<td>35.80 (11.45)</td>
<td>36.00 (11.76)</td>
<td>35.90 (9.58)</td>
</tr>
</tbody>
</table>

3.1.1 Results of Three-Way ANOVA

A three-way mixed ANOVA (A: groups: control/experimental 1-3, B: time: pre/post, C: words: content/function) was conducted and the results are shown in Table 3. Figure 1 shows the means of accurate lexical recognition in percentage for each two of the three factors.

![Figure 1](image)

Figure 1. Words successfully identified in percentage for each two of the three factors. (**p < .01)
The three-way interaction was not significant \(F(3, 202) = 0.171, p = .916 \text{ ns, } \eta^2 = .000\). Neither was there significant interaction between groups and words \(F(3, 202) = 1.540, p = .205 \text{ ns, } \eta^2 = .002\). However, significant interactions were found between groups and time \(F(3, 202) = 6.854, p = .000 < .001, \eta^2 = .006\) as well as time and words \(F(1, 202) = 342.838, p = .000 < .001, \eta^2 = .97\), so that respective simple main effects were examined. Differences in mean values at the posttest between four groups were assessed with multiple comparison procedure using Tukey-Kramer method.

As for the interaction between groups and time, although the simple main effects of time in all the four groups were significant (Control Group: \(F(1, 202) = 16.755, p = .000 < .001, \eta^2 = .043\), Experimental Group 1: \(F(1, 202) = 21.751, p = .000 < .001, \eta^2 = .056\), Experimental Group 2: \(F(1, 202) = 95.763, p = .000 < .001, \eta^2 = .246\), Experimental Group 3: \(F(1, 202) = 53.001, p = .000 < .001, \eta^2 = .136\), those of groups were not significant in the pretest \(F(3, 404) = 0.844, p = .470 \text{ ns, } \eta^2 = .006\), while in the posttest they were significant \(F(3, 404) = 3.664, p = .013 < .05, \eta^2 = .026\). This means that, even though the participants in all the four groups fared better in the posttest than in the pretest, difference in word recognition, which did not exist between the four groups in the pretest, emerged in the posttest after the treatment.

In addition, the results of multiple comparison procedure indicated that the participants in Experimental Group 2, who constantly listened to the textbook's CD at 1.5 times the original rate, fared significantly better in the posttest than the other three groups (Control Group & Experimental Group 2: \(t = 3.513, p = .003 < .01\), Experimental Groups 1 & 2: \(t = 4.886, p = .000 < .001\), Experimental Groups 2 & 3: \(t = 3.278, p = .006 < .01\)). Although, as the graph shows, Experimental Group 3's word recognition also improved better than those of Control Group and Experimental Group 1, the difference in the posttest was not significant (Control Group & Experimental Group 3: \(t = 0.156, p = .999 \text{ ns, Experimental Groups 1 & 3: } t = 1.358, p = .520 \text{ ns}\).
As for the interaction between time and words, although the simple main effects of time for content words were significant \((F(1, 404) = 488.907, \ p = .000 < .001, \ η_p^2 = .544)\), those for function words were not significant \((F(1, 404) = 1.979, \ p = .160 \ n.s., \ η_p^2 = .007)\) and, although the simple main effects of words in the pretest was not significant \((F(1, 404) = 0.221, \ p = .638 \ n.s., \ η_p^2 = .000)\), those in the posttest was significant \((F(1, 404) = 614.241, \ p = .000 < .001, \ η_p^2 = .603)\). These results mean that recognition of content words improved across the groups, while that of function words did not. As a result, difference in recognition between content and function words, which did not exist in the pretest, emerged in the posttest.

### 3.1.2 Results of Two-Way ANOVAs

In order to more thoroughly analyze the effects of the treatment on word recognition, a two-way mixed ANOVAs (A: groups: control/experimental 1-3, B: time: pre/post) were conducted separately for content and function words. As a result, significant interactions were found for both content \((F(3, 202) = 4.636, \ p = .004 < .01, \ η_p^2 = .007)\) and function \((F(3, 202) = 4.654, \ p = .000 < .001, \ η_p^2 = .012)\) words, so that simple main effects were examined. Figure 2 shows the means of accurate content and function word recognition in percentage by each group at the pretest and the posttest.

![Content and function words successfully identified in percentage at the pretest and the posttest. (***p < .01, **p < .05)](image)

First, recognition of content words improved significantly across all the four groups between the pretest and the posttest (Control Group: \(F(1, 202) = 76.015, \ p = .000 < .001, \ η_p^2 = .392\), Experimental Group 1: \(F(1, 202) = 93.190, \ p = .000 < .001, \ η_p^2 = .113\), Experimental Group 2: \(F(1, 202) = 329.362, \ p = .000 < .001, \ η_p^2 = .401\), Experimental Group 3: \(F(1, 202) = 121.718, \ p = .000 < .001, \ η_p^2 = .148\). That of function words, however, deteriorated significantly for Control Group \((F(1, 202) = 8.219, \ p = .005 < .01, \ η_p^2 = .037)\) and Experimental Group 1 \((F(1,
Experimental Group 2 fared significantly better than the other three groups, while the differences remained unchanged.

Second, difference between the groups, which was nonexistent at the pretest (content: \( F(3, 404) = 1.386, p = .247 \), function: \( F(3, 404) = 0.381, p = .767 \)), was made after the treatment at the posttest (content: \( F(3, 404) = 5.218, p = .002 < .01, \eta_p^2 = .054 \), function: \( F(3, 404) = 3.117, p = .027 < .05, \eta_p^2 = .032 \)) for both content and function word recognition. Though the difference between the groups at the posttest was greater in recognition of content words than in that of function words, the results of multiple comparison procedure, using Tukey-Kramer method, illustrate that, in both content and function word recognition, Experimental Group 2 fared significantly better than the other three groups, while the differences between the other three groups were not significant (Table 4).

### Table 4
The Results of Multiple Comparison for Content and Function Word Recognition at the Posttest (n = 206)

<table>
<thead>
<tr>
<th>Group 1 (I)</th>
<th>Group 2 (J)</th>
<th>Content-Word Recognition</th>
<th>Function-Word Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Difference (J-I)</td>
<td>t</td>
</tr>
<tr>
<td>Control</td>
<td>Exp 1</td>
<td>-2.887</td>
<td>1.370</td>
</tr>
<tr>
<td>Control</td>
<td>Exp 2</td>
<td>6.477</td>
<td>3.562</td>
</tr>
<tr>
<td>Control</td>
<td>Exp 3</td>
<td>0.581</td>
<td>0.274</td>
</tr>
<tr>
<td>Exp 1</td>
<td>Exp 2</td>
<td>9.365</td>
<td>5.109</td>
</tr>
<tr>
<td>Exp 1</td>
<td>Exp 3</td>
<td>3.468</td>
<td>1.625</td>
</tr>
<tr>
<td>Exp 2</td>
<td>Exp 3</td>
<td>-5.896</td>
<td>3.190</td>
</tr>
</tbody>
</table>

** ***: \( p < .001 \), **: \( p < .01 \), *: \( p < .05 \)

### 3.1.3 Results of Chi-Square Tests
In order to examine the effects of treatment in detail, chi-square tests were conducted. The numbers of right (R) and wrong (W) transcriptions by the three experimental and control groups for all the 50 words in the posttest were computed and in which of the 50 words in the posttest there were significant differences in recognition between the four groups was examined.

Results are shown in Table 5. In terms of chi-square values, a significant difference was found in 8 words. In addition, significant standardized residuals were found in 4 other words for Experimental Group 2, and in still another word for Experimental Group 3, even though the chi-square value was not significant. Among these 13 words, in which a significant difference in recognition was found between the groups, participants in Experimental Group 2 fared significantly better than the other three groups in 11 words, seven content and four function words. The greatest difference was found in the word No. 30 yet \( (\chi^2 (3) = 18.991, p = .000 < .001, \text{Cramer’s V} = .304) \).
Table 5
Words in Which There Was Significant Difference in Recognition Between the Control and Three Experimental Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>2 popular</th>
<th>3 such</th>
<th>4 should</th>
<th>6 math</th>
<th>10 aren’t</th>
<th>11 where</th>
<th>17 giving</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (n = 42)</td>
<td>39</td>
<td>18</td>
<td>24</td>
<td>34</td>
<td>8</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>Exp 1 (n = 43)</td>
<td>35</td>
<td>19</td>
<td>22</td>
<td>27</td>
<td>14</td>
<td>17</td>
<td>24</td>
</tr>
<tr>
<td>Exp 2 (n = 83)</td>
<td>81</td>
<td>2</td>
<td>56</td>
<td>27</td>
<td>75</td>
<td>8</td>
<td>65</td>
</tr>
<tr>
<td>Exp 3 (n = 40)</td>
<td>33</td>
<td>7</td>
<td>22</td>
<td>18</td>
<td>37</td>
<td>3</td>
<td>27</td>
</tr>
</tbody>
</table>

Con: 0.41 0.41 1.90 1.90 0.60 0.60 0.02 0.02 1.04 1.04 1.72 1.72 0.55 0.55

Residuals: Exp 1 1.49 1.49 1.37 1.37 3.54 3.54 ** 3.80 3.80 ** 3.98 3.98 ** 2.38 2.38 ** 0.60 0.60 ** 1.43 1.43

Cramer’s V = .304.

The greatest difference was found in the word No. 30 yet giving.

In 8 words, significant standardized residuals were found in 4 other words for the Control Group. In total, significant differences were found in 8 words. In addition, significant standardized residuals were found in 4 other words for the Control Group.

In order to examine the effects of treatment in detail, chi-square tests were conducted. The results of the ANOVA shows that the interaction was not significant ($F(3, 403) = 0.381$, $p = .573$).

Second, difference between the groups, which was nonexistent at the pretest (content: $\eta^2 = 0.027$, function: $\eta^2 = 0.005$), was significant ($F(3, 403) = 5.218$, $p = .005$).

The results of the ANOVA shows that the interaction was significant ($F(3, 403) = 9.439$, $\eta^2 = 0.143$).

The difference between the groups at the posttest was greater in recognition of content words than in that of function words, the results of multiple comparison procedure, using Tukey-Kramer method, illustrate that, in both content and function word recognition. Though the difference between the groups at the posttest was greater in recognition of content words than in that of function words, the results of multiple comparison procedure, using Tukey-Kramer method, illustrate that, in both content and function word recognition. Though the difference between the groups at the posttest was greater in recognition of content words than in that of function words, the results of multiple comparison procedure, using Tukey-Kramer method, illustrate that, in both content and function word recognition. Though the difference between the groups at the posttest was greater in recognition of content words than in that of function words, the results of multiple comparison procedure, using Tukey-Kramer method, illustrate that, in both content and function word recognition.

3.2 Listening Comprehension Tests

Table 6 shows the descriptive statistics of the pretest (Cronbach’s alpha = .783) and the posttest (Cronbach’s alpha = .761). Based on this data, a two-way mixed ANOVA (A: groups: control/experimental 1-3, B: time: pre/post) was conducted.

The results of the ANOVA shows that the interaction was not significant ($F(3, 202) = 1.950$, $p = .123$). Yet, even though the main effect of time was significant ($F(1, 202) = 298.346$, $p = .000 < .001$, $\eta^2 = .316$), that of groups was not significant ($F(3, 202) = 0.122$, $p = .947$).
These results mean that, although there was improvement in listening comprehension across all the groups from the pretest to the posttest, there was no effect of difference in the speech rate at which the participants listened during the treatment on listening comprehension.

4. Discussion

First, as to RQ1, whether steady exposure to a faster rate of speech in class is effective in improving word recognition by Japanese EFL listeners with lower levels of proficiency, the results of the present study showed that half-a-year treatment in which the participants were given listening materials at different speech rates was effective on their lexical recognition. In addition, as to RQ2, whether there is any specific speech rate at which low-proficiency listeners fare best in word recognition, the results showed that it was most effective at the compressed rate of about 67 percent, that is, 1.5 times the baseline speech rate and that any other compressed rate was found not to have any significant effects. Finally, as to RQ3, whether the training practice has any positive effect on listening comprehension, the participants in all the groups did enhance their listening comprehension after the treatment. However, no difference in comprehension was found between the groups who constantly listened at different speech rates throughout the treatment.

These results suggest that, when L2 learners, especially low-proficiency learners, are trained by constantly listening to English at the ‘right’ compressed rate, there seem to be some positive effects on their recognition of words. The results of the present study show that the ‘right’ rate is 1.5 times the original rate. In addition, regardless of word categories, content or function, Experimental Group 2 fared significantly better in the posttest than the other three groups. However, the gain in function word recognition that Experimental Group 2 made between the two tests was smaller than that in content word recognition and, for all the other groups, recognition of function words deteriorated in the posttest. This may be because it takes more than improvement in bottom-up processing to successfully recognize function words, such as intervention from the top-down processing.

Nevertheless, in the analysis of word by word, the word in which the greatest difference was found in the posttest was a function word. It may safely be said, therefore, that the speech rate that is effective in training elementary-level Japanese EFL learners to improve word recognition of the normal speech rate (around 130 wpm) is 1.5 times the baseline rate (around 190 wpm), even though what constitutes the normal speech rate for L2 learners is still another issue that should be discussed.

On the other hand, as far as the comprehension is concerned, no effect of different speech rates during the treatment was found, even though their comprehension improved after half-a-year training sessions. There are some possibilities about this. First, in listening, semantic and syntactical processing must be done in real time, as the learners listen to the passage. This means that higher speech rates require them to process the incoming speech at a higher rate, forcing them...
to make adjustments in order to keep up with the heightened rate. During the treatment sessions, it is improbable that this might have happened, which is why difference in speech rates during the treatment did not translate to difference in comprehension and Control Group fared as well as the three experimental groups.

According to Adank and Devlin (2010), semantic and syntactical processing takes place in the left hemisphere and prosodic and acoustic processing in the right hemisphere. They hold that, after L1 speakers listened to highly compressed speech about double the rate over several sessions, adaptation-related changes were observed both in the left and right hemisphere, which indicated that they made adjustments at both linguistic and acoustic levels. During the high-speed listening sessions, not only were L1 listeners able to make adjustments in getting used to the higher speech rate acoustically, they were also able to comprehend the speech at a higher rate. This suggests that the treatment in the present study may be effective even in comprehension, in the case of L1 speakers or possibly L2 learners of near-native proficiency. As the results of the present study shows, however, treatment of a kind in which high-speed listening is constantly given is effective in improving acoustic processing but not linguistic processing, in the case of low-proficiency L2 learners.

Another reason why the treatment did not make difference in the participants’ comprehension may be that speech rates consist of two main factors: pause duration and articulation rates and that it is the length of pauses that affects listeners’ comprehension and not the rates at which the speech is articulated (Kano & Saito, 1997; Sugai, Yamane, & Kanzaki, 2016). Kano and Saito (1997) and Sugai et al. (2016) both suggest that the more influential parameter in listening in terms of word recognition is articulation rates and that the one that affects comprehension is pauses inserted at grammatical junctures. Although, as Kano and Saito (1997) suggests, improved word recognition leads to improved comprehension, it is possible that pauses play a more influential role, which is supported by Suzuki (1991) and Sugai, Kanzaki, and Yamane (2007).

Finally, as far as the group-by-group improvement in word recognition after the treatment, Experimental Group 3, who constantly listened to the textbook’s CD at double the original rate, fared second best after the treatment, even though the difference was not significant. According to Dupoux and Green (1997), more highly compressed stimuli require more time for improvement than less compressed ones and performance improves with increased exposure to compressed speech. This suggests that the results obtained for the group who listened during the treatment at double the baseline rate may have turned significant if the span of the treatment had been longer.

5. Conclusion

The present study showed that, if elementary-level L2 learners continue to listen at a higher speech rate, they make improvements in word recognition of the baseline speech rate and that the
most effective rate is 1.5 times the original rate. However, this does not affect comprehension. Nevertheless, the importance of word recognition cannot be neglected because, once L2 learners make wrong guesses in recognizing words, they are often unable to modify those previous guesses as they listen, which will lead to misunderstandings of the whole context (Yamauchi, Yamato, & Kida, 2016).

There are, however, limitations to this study. First, the participants in the present study were at an elementary-proficiency level. The results would have been different with higher-proficiency learners. Further study should explore the relationship between learners’ proficiency and the effects of high-speed listening sessions. Second, in the present study, compressed speech rates in the training sessions were arbitrarily set at 1.2, 1.5, and 2.0 times the original rate. Even though the group who listened at 1.5 times fared best, whether a group of 1.4, 1.6, or 1.7 times the original rate might have performed better is unknown. Despite these limitations, it is hoped that the present study becomes a stepping stone for further study on the effects of speech rate changes on L2 learners’ word recognition and listening comprehension.

References


Appendix: The Scripts of the Word Recognition Posttest (Words Blanked Out for Transcription Are in Bold Type)

[Dialogue]
1. Wow! There are so many people (1)here, Maiko. / Yeah, this is one of the most (2)popular beaches in Japan. / I’ve never seen (3)such a crowded beach before. I (4)should take a picture to show to my family in Canada. / I’ll take a picture with you (5)in it. Stand over there.
3. Excuse me. (11)Where is the cafeteria at this university? / There are two of them. The one (12)in Smith Hall is better. / Great. Where is that? / Do you see (13)that tall building over there? That’s it. You’ll (14)see the cafeteria as soon as you enter it.
4. Did you sleep (15)well last night, Pam? / Yeah, Dad. I dreamed I was a (16)famous singer. I was (17)giving a concert. / Well, (18)you’re good at singing. Maybe your dream will come (19)true. / That would be nice. But I’ll need to take singing lessons (20)first.
5. Clarksville Public Library. / Hi. I think I (21)left my coat on a chair when I was (22)there this morning. / What floor were you (23)on, sir? / The (24)third floor. I was sitting (25)near the newspaper section.

[Monologue]
6. Dr. Tanaka is a (26)dentist in Tokyo. Recently, many foreign (27)patients have been coming to his clinic. (28)Some of them can’t speak Japanese, and Dr. Tanaka wants to communicate with them in English. As a (29)result, he has started taking English lessons. Dr. Tanaka can’t speak very well (30)yet, but he is studying hard.
7. Nicole bought a (31)used car last month. (32)While she was driving it yesterday, she noticed something strange. Whenever she (33)turned a corner, the car made a loud noise. But when she drove (34)straight, the car sounded normal. Nicole (35)plans to take it to a repair shop this weekend.
8. Tim’s family moved to a (36)different town this week. Tim will (37)start going to his new high school next Monday. Before then, he has to do some work (38)that his new teachers gave him to help him catch (39)up with his classmates. Tim is (40)looking forward to making new friends at the school.
9. Attention, ladies and gentlemen. The Springfield Summer Parade will begin shortly. I’d like to (41)ask everyone standing in the street (42)to take a few steps back. Please move onto the sidewalk so (43)that the people marching in the parade can get (44)through. Thank you for your cooperation.
10. Today, many people (45)around the world enjoy skateboarding. Skateboards became (46)popular in California in the 1960s. At the time, there were many people in southern California (47)who went surfing. But (48)when the waves at the beach were not big enough for surfing, surfers looked (49)for activities to do on land. They started to buy skateboards and ride (50)them on sidewalks and roads.