Comparative study of intensity in the speech of native speakers and Japanese speakers of English

Tomoko Nariai and Kazuyo Tanaka

1 Department of Career Development and Liberal Arts, Tokiwa Junior College, Miwa 1–430–1, Mito, 310–8585 Japan
2 Faculty of Library, Information and Media Science, University of Tsukuba, Kasuga 1–2, Tsukuba, 305–8550 Japan

(Received 29 August 2012, Accepted for publication 18 June 2013)

Abstract: This paper presents a comparative analysis of the intensity of words in sentences uttered by Japanese speakers of English and native speakers of English (Japanese English and native English, henceforth). We investigate two parameters: intensity, which is defined here as the integral of the power for each word, and power peak, which is defined as the peak of the power for each word. The analyses reveal differences in both word class and word position. For word class, the nouns, interrogatives and negatives for Japanese English are produced with less intensity, whereas most function words are produced with more intensity than are their native English counterparts. For word position, Japanese English sentence-final words are produced with less intensity than are their native English counterparts. Also, Japanese English sentence-initial words show lower power peak, and the sentence-final words show higher power peak than are their native English counterparts. Detailed analyses reveal a correlation between the above results and subjects’ English proficiency. The results for word class can be explained as a result of the Japanese speakers inserting Japanese focus into English utterances. The results for word position are explained such that while sentence-initial strengthening does not affect sentence-initial power peak in Japanese English, an irregularity of final lengthening affects sentence-final intensity.

Keywords: Power, Prosody, Second language learning, Speech analysis, CALL

PACS number: 43.70.Kv, 43.70.Fq [doi:10.1250/ast.35.42]

1. INTRODUCTION

Research in language production has revealed that prosodic features play a pivotal role in human communication [1]. Therefore, learning prosodic patterns is essential to the acquisition of second-language pronunciation.

There has been a significant increase in the amount of research that examines the characteristics of Japanese English. Despite the quantity and quality of studies addressing this topic in the past two decades, two main areas remain largely unexamined.

Recent lines of research have focused on the acoustic analyses of the segmental level [2–4], while research into the prosodic level remains relatively sparse. In addition, few studies of Japanese English have involved the quantitative examination of intensity. These characteristics, when revealed by speech analysis, can be used to improve foreign language teaching and to construct more extensible language learning systems.

Second, most-mediated learning tools, such as computer-assisted language learning (CALL), have shown their ability to improve learners’ production of phonemes. In contrast, little progress has been made in using CALL to teach the categorical nature of prosody. In such situations, shadowing is implemented as one of the most effective methods for helping learners to acquire English speaking and listening skills [5].

Shadowing, as an English discipline system, encourages learners to master natural-sounding English by shadowing native speakers’ speech, so it is a useful tool for learners seeking mastery of English. However, shadowing encourages learners to speak English in a stream-like manner, often at speeds that are too fast for learners to keep up with.

In this paper, the intensity of Japanese English words in sentence utterances is compared to the intensity of the same words spoken by native English. The words are then analyzed by word class and sentential position.
Intensity is defined in subsection 2.4 as the integral of the power for each word. For a detailed consideration of intensity, we investigate the power peak, which is also defined in subsection 2.4 as the peak of the power for each word.

The reason for investigating intensity is based on the linguistic aspect of prosody in both English and Japanese. For English, stress accent is produced by a higher pitch, a longer duration and a greater intensity [6]. For Japanese, pitch accent is produced by a change in pitch [7]. Second, previous studies of the perception of English stress have revealed that English speakers are sensitive to pitch, duration and, to a lesser extent, amplitude. In contrast, Japanese speakers are sensitive almost exclusively to pitch [8]. These prosodic differences between English and Japanese presumably shape the characteristics of intensity that appear in the combination of power and duration in Japanese English.

Section 2 presents sample data and the analysis methods. Section 3 presents the analytical results, including the results of analyses after subjects were classified according to English proficiency. A discussion is presented in Sect. 4, and the conclusion is given in Sect. 5.

2. SPEECH SAMPLES USED IN THE ANALYSIS

2.1. Sample Sentences
The sentence text set of this speech dataset was the same as that of the TIMIT dataset [9]. In the analysis, 100 sentences were chosen; the TIMIT sentence numbers were 001–030, 211–260, and 441–460. There were 707 words in total.

2.2. Subjects
2.2.1. English speakers
The group of native English speakers consisted of 10 subjects: five males and five females, aged between 20 and 40 years. Subjects were English teachers living in Japan, and were from the United Kingdom, Canada, New Zealand, Australia, and the United States.

2.2.2. Japanese speakers
The group of native Japanese speakers consisted of 24 subjects: 12 males and 12 females, aged between 20 and 30 years.

2.3. Recording Conditions
The recording equipment comprised a SONY F 780 dynamic microphone, and a DAT SONY 2000ES recording device. The distance from the mouth to the microphone was kept constant at about 30 cm.

The subjects were given sufficient time to practice reading the speech materials before recording. Subjects were asked to enunciate clearly and to utter the sentence repeatedly until the speech sample was recorded properly. No other specific instructions were given for producing English utterances.

The 10 English subjects uttered 100 sentences each. The group of 12 Japanese subjects uttered 50 sentences each. The TIMIT sentence numbers of these sentences were 001–030 and 211–230. A second group of 12 Japanese subjects uttered the remaining 50 sentences.

Each sentence utterance was sampled at a rate of 48 kHz and quantized into 16 bits.

2.4. Analysis Method
The intensity value of a word was calculated from the power pattern of each utterance, using WaveSurfer/Snack Sound Toolkit [10]. The power pattern was extracted from the speech wave with a Hamming window of length 20 ms, at a frame interval of 10 ms, and converted into dB. The maximum of the power value was 80 dB, and the minimum was 0 dB. These were the default settings in WaveSurfer. The power value sequence was manually segmented into a word sequence, as in Fig. 1.

We investigate two parameters: intensity and power peak. These are defined as follows.

Intensity of word (i) = Sum of the power values for the section of word i.

Power peak of word (i) = Maximal value of the power values for the section of word i.

[Equations for intensity and power peak are provided here.]

The prosodic values of individual words were then normalized by the average of those values contained in the sentence. The prosody of word i is represented by the following.

$x_j(i)' = x_j(i)/\bar{x}_j$ (1)  
$y_j(i)' = y_j(i)/\bar{y}_j$ (2)

where,

$\bar{x}_j = \frac{\sum_{i=1}^{L} x_j(i)}{L}$  
$\bar{y}_j = \frac{\sum_{i=1}^{L} y_j(i)}{L}$

$x_j(i)$: the intensity of word i uttered by English subject j 
$y_j(i)$: the intensity of word i uttered by Japanese subject j 
$L$: number of words in the corresponding sentence
In this study, an index, denoted by $R$, is used for representing the difference between the two groups. This $R$ is conventionally used in statistical pattern recognition and is the ratio of the between-group variance to the within-group variance, known as Fisher’s ratio in linear discriminant analysis. A large $R$ indicates a significant difference between the two groups. $R$ is calculated as follows.

$$ R(i) = \frac{(\bar{x}(i) - \bar{y}(i))^2}{(\sigma_x(i) + \sigma_y(i))^2} \quad (3) $$

where,

$$ \bar{x}(i) = \frac{1}{N} \sum_{j=1}^{N} x_j(i) \quad (4) $$

$$ \bar{y}(i) = \frac{1}{M} \sum_{j=1}^{M} y_j(i) \quad (5) $$

$$ \sigma_x(i) = \frac{1}{N} \sum_{j=1}^{N} (x_j(i) - \bar{x}(i))^2 \quad (6) $$

$$ \sigma_y(i) = \frac{1}{M} \sum_{j=1}^{M} (y_j(i) - \bar{y}(i))^2 \quad (7) $$

$N$: number of English subjects

$M$: number of Japanese subjects

$R$ values were calculated for individual words, then, we obtained $R_1$ to $R_{707}$.

In the analysis, “ntv>jpe” indicates that a word for English subjects “ntv” has higher intensity than that for Japanese subjects “jpe,” and “ntv<jpe” indicates the reverse. In the present study, only cases of significant difference between the two groups are counted; specifically, the inequality, “ntv>jpe” or “ntv<jpe,” holds only for cases that satisfy $R > 0.2$.

2.5. Classification

2.5.1. Word class

Words were classified into content words or function words in accordance with their grammatical functions [11].

Content words were further classified into nouns (core of noun phrases), verbs (core of verb phrases), adjectives (playing the role of adjective in sentences) and adverbs (playing the role of adverb in sentences).

In the following tables, the number of nouns, represented by ‘noun,’ totals 190 words; adjectives, ‘adj,’ total 132; verbs, ‘verb,’ total 106; adverbs, ‘adv,’ total 24. The sum of content words, ‘C-sum,’ totals 452.

Function words were also further classified into five groups: /interrogative, negative/, /conjunction, preposition/, /be, auxiliary verb, do/, article and pronoun.

In the following tables, /interrogative, negative/ words, represented by ‘int, ng’ total 9; /conjunction, preposition/ words, ‘cnj, prp,’ total 84; /be, auxiliary verb, do/ words, ‘be,’ total 37; articles, ‘art,’ total 76; pronouns, ‘pron,’ total 49. The sum of function words, ‘F-sum,’ total 255.

2.5.2. Word position

Words were divided into three positions: words at the beginning of sentences (sentence-initial); words within sentences (within); words at the ends of sentences (sentence-final).

In the following tables, sentence-initial words, represented by ‘initial,’ total 100; ‘within’ words total 507; sentence-final words, represented by ‘final,’ total 100.

3. RESULTS

3.1. Word Class

3.1.1. Word intensity

Table 1 shows the intensity results for content words. The columns in the table indicate word class. The second row from the top indicates the number of words classified into the word class. The third row indicates the number of words that satisfy $R > 0.2$. The fourth and fifth rows indicate the number of words that satisfy “ntv>jpe” and “ntv<jpe,” respectively.

Out of 190 nouns, 111 satisfy $R > 0.2$, 87% of which satisfy “ntv>jpe.” Out of 132 adjectives, 61 satisfy $R > 0.2$, 53% of which satisfy “ntv<jpe.” Out of 24 adverbs, 11 satisfy $R > 0.2$, 64% of which satisfy “ntv<jpe.” For all 452 content words, 234 satisfy $R > 0.2$, 69% of which satisfy “ntv<jpe.”

The results suggest that more than half the nouns and adjectives satisfy “ntv>jpe,” and more than half the verbs and adverbs satisfy “ntv<jpe.” Therefore, we can identify the following characteristics of intensity for content words in Japanese English.

Figure 2 shows an example taken from TIMIT259: “Do they allow atheists in church?” Satisfying “ntv>jpe” is the nouns, atheists($0.73$) and church($0.97$), where the $R$ is given in the parenthesis. In contrast, the verb, allow($0.39$), satisfies “ntv<jpe.”

Figure 3 shows an example taken from TIMIT451: “The toddler found a clamshell near the camp site.” The nouns and adjective, toddler($0.26$), clamshell($0.95$), camp($0.73$)

<table>
<thead>
<tr>
<th>Table 1 Results for content words.</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>190</td>
</tr>
<tr>
<td>97</td>
</tr>
<tr>
<td>(87%)</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>(53%)</td>
</tr>
</tbody>
</table>
and site (0.81), satisfy “ntv>jpe.” In contrast, the verb, found (0.44), satisfies “ntv<jpe.”

Among the above, the probability of “ntv>jpe” for nouns is high. This suggests that Japanese English nouns are produced with less intensity than those by native English.

Table 2 shows the intensity results for function words.

Table 2: Results for function words.

<table>
<thead>
<tr>
<th></th>
<th>int, ng</th>
<th>cnj</th>
<th>be</th>
<th>art</th>
<th>prn</th>
<th>F-sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of words</td>
<td>9</td>
<td>84</td>
<td>37</td>
<td>76</td>
<td>49</td>
<td>255</td>
</tr>
<tr>
<td>( R &gt; 0.2 )</td>
<td>4</td>
<td>71</td>
<td>25</td>
<td>48</td>
<td>27</td>
<td>175</td>
</tr>
<tr>
<td>( ntv&gt;jpe )</td>
<td>4 (100%)</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>( ntv&lt;jpe )</td>
<td>71 (100%)</td>
<td>21</td>
<td>45</td>
<td>22</td>
<td>159 (84%)</td>
<td>22 (94%)</td>
</tr>
</tbody>
</table>

and site (0.81), satisfy “ntv>jpe.” In contrast, the verb, found (0.44), satisfies “ntv<jpe.”

Among the above, the probability of “ntv>jpe” for nouns is high. This suggests that Japanese English nouns are produced with less intensity than those by native English.

Table 2 shows the intensity results for function words.

Out of 11 /interrogative, negative/ words, 4 satisfy \( R > 0.2 \), all of which satisfy “ntv>jpe.” In contrast, over 80% of the words for /conjunction, preposition/, /be/, auxiliary verb, do/, article, and pronoun satisfy “ntv<jpe.” From these results, we can identify the following characteristics of intensity for function words in Japanese English.

Figure 4 shows an example taken from TIMIT221: “How permanent are their records?” The interrogative, How (0.39), satisfies “ntv>jpe.” In contrast, the be, are (4.64), and the pronoun, their (1.26), satisfy “ntv<jpe.”

Figure 5 shows an example taken from TIMIT242: “The courier was a dwarf,” where the articles, The (0.73) and a (0.21), and the be, was (2.97), satisfy “ntv<jpe.”

Considering the results of the sum of function words that 91% of the 175 words, that satisfied \( R > 0.2 \), satisfy “ntv<jpe,” Japanese English function words are produced with greater intensity than those by native English.

3.1.2. Power peak for word

Table 3 shows the power results for content words. The results suggest that more than half the nouns, adjectives, verbs and adverbs satisfy “ntv>jpe.”

Table 4 shows the power results for function words. The results suggest that /interrogative, negative/ satisfy “ntv>jpe,” whereas more than half the /conjunction, preposition/ and /be, auxiliary verb, do/, articles and pronouns satisfy “ntv<jpe.”

From these results, we can identify the following characteristics of the power peak in Japanese English.
Figure 6 shows an example in TIMIT29: ‘‘Help Greg to pick a peck of potatoes’’, where the verbs, Help\(0.65\) and pick\(0.93\), and the nouns, Greg\(0.64\) and potatoes\(0.21\), satisfy ‘‘ntv\(>\)jpe.’’ In contrast, the prepositions, to\(0.58\) and of\(1.71\), and the article, a\(1.36\), satisfy ‘‘ntv\(<\)jpe.’’

Of all others, the probability of ‘‘ntv\(>\)jpe’’ in adverbs and /interrogative, negative/, whereas that of ‘‘ntv\(<\)jpe’’ in /conjunction, preposition/ are high. These signify low power for adverbs and /interrogative, negative/, whereas high power for /conjunction, preposition/ in Japanese English.

In addition, comparison between the results of sum for content words and function words reveals low power for content words, whereas high power for function words in Japanese English.

### Table 3 Power results for content words.

<table>
<thead>
<tr>
<th></th>
<th>noun</th>
<th>adj</th>
<th>verb</th>
<th>adv</th>
<th>C-sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of words</td>
<td>190</td>
<td>132</td>
<td>106</td>
<td>24</td>
<td>452</td>
</tr>
<tr>
<td>(R &gt; 0.2)</td>
<td>77</td>
<td>58</td>
<td>34</td>
<td>9</td>
<td>178</td>
</tr>
<tr>
<td>ntv(&gt;)jpe</td>
<td>45 (58%)</td>
<td>42 (72%)</td>
<td>23 (68%)</td>
<td>7 (78%)</td>
<td>117</td>
</tr>
<tr>
<td>ntv(&lt;)jpe</td>
<td>32 (62%)</td>
<td>16 (51%)</td>
<td>11 (69%)</td>
<td>2 (71%)</td>
<td>61</td>
</tr>
</tbody>
</table>

### Table 4 Power results for function words.

<table>
<thead>
<tr>
<th></th>
<th>int, ng</th>
<th>cnj</th>
<th>be</th>
<th>art</th>
<th>prn</th>
<th>F-sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of words</td>
<td>9</td>
<td>84</td>
<td>37</td>
<td>76</td>
<td>49</td>
<td>255</td>
</tr>
<tr>
<td>(R &gt; 0.2)</td>
<td>3</td>
<td>49</td>
<td>18</td>
<td>18</td>
<td>23</td>
<td>111</td>
</tr>
<tr>
<td>ntv(&gt;)jpe</td>
<td>3 (100%)</td>
<td>4 (100%)</td>
<td>7 (100%)</td>
<td>6 (100%)</td>
<td>10 (100%)</td>
<td>30</td>
</tr>
<tr>
<td>ntv(&lt;)jpe</td>
<td>45 (92%)</td>
<td>11 (22%)</td>
<td>12 (22%)</td>
<td>13 (22%)</td>
<td>81 (22%)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5 Intensity results for word position.

<table>
<thead>
<tr>
<th></th>
<th>initial</th>
<th>within</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of words</td>
<td>100</td>
<td>507</td>
<td>58</td>
</tr>
<tr>
<td>(R &gt; 0.2)</td>
<td>56</td>
<td>295</td>
<td>52 (90%)</td>
</tr>
<tr>
<td>ntv(&gt;)jpe</td>
<td>22 (45%)</td>
<td>103 (45%)</td>
<td>52 (90%)</td>
</tr>
<tr>
<td>ntv(&lt;)jpe</td>
<td>34 (61%)</td>
<td>192 (65%)</td>
<td>6 (10%)</td>
</tr>
</tbody>
</table>

Figure 6 shows an example in TIMIT29: ‘‘Help Greg to pick a peck of potatoes,’’ where the verbs, Help\(0.65\) and pick\(0.93\), and the nouns, Greg\(0.64\) and potatoes\(0.21\), satisfy ‘‘ntv\(>\)jpe.’’ In contrast, the prepositions, to\(0.58\) and of\(1.71\), and the article, a\(1.36\), satisfy ‘‘ntv\(<\)jpe.’’

3.2. Word Position

3.2.1. Word intensity

Table 5 shows the intensity results for word position.

The results suggest that the sentence-initial and within words satisfy ‘‘ntv\(<\)jpe,’’ whereas the sentence-final words satisfy ‘‘ntv\(>\)jpe.’’ Among these, the probability of ‘‘ntv\(>\)jpe’’ for sentence-final words is high. This means that the initial and within positions of sentences do not affect ‘‘ntv\(>\)jpe’’ or ‘‘ntv\(<\)jpe,’’ whereas the end position results in ‘‘ntv\(>\)jpe.’’

Figure 7 shows an example taken from TIMIT13: ‘‘Swing your arms as high as you can.’’ The sentence-initial verb, Swing\(0.21\), satisfies ‘‘ntv\(<\)jpe.’’ The within noun and adjective, arm\(0.87\) and high\(0.23\), satisfy ‘‘ntv\(>\)jpe.’’ The within function words, as\(0.61\), as\(0.46\) and you\(2.88\), satisfy ‘‘ntv\(<\)jpe,’’ whereas the sentence-final function word, can\(0.66\), satisfies ‘‘ntv\(>\)jpe.’’

This shows that Japanese English sentence-final words are produced with less intensity than those by native English.

3.2.2. Power peak for word

Table 6 shows the power results for word position. The results suggest that the sentence-initial words satisfy ‘‘ntv\(>\)jpe,’’ whereas the within and sentence-final words satisfy ‘‘ntv\(<\)jpe.’’ In particular, the probabilities of ‘‘ntv\(>\)jpe’’ for sentence-initial words and ‘‘ntv\(<\)jpe’’ for sentence-final words are high. These signify that the within

### Table 6 Power results for word position.

<table>
<thead>
<tr>
<th></th>
<th>initial</th>
<th>within</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of words</td>
<td>100</td>
<td>507</td>
<td>58</td>
</tr>
<tr>
<td>(R &gt; 0.2)</td>
<td>56</td>
<td>295</td>
<td>52 (90%)</td>
</tr>
<tr>
<td>ntv(&gt;)jpe</td>
<td>22 (45%)</td>
<td>103 (45%)</td>
<td>52 (90%)</td>
</tr>
<tr>
<td>ntv(&lt;)jpe</td>
<td>34 (61%)</td>
<td>192 (65%)</td>
<td>6 (10%)</td>
</tr>
</tbody>
</table>
position of sentences does not affect “ntv>jpe” or “ntv<jpe.” In contrast, the initial position of sentences results in “ntv>jpe,” whereas the end position results in “ntv<jpe.”

Figure 8 shows the results for TIMIT448: “Don’t look for cheap valuables in a bank vault.” The sentence-initial auxiliary verb, *Do*$_{(1.75)}$ satisfies “ntv>jpe.” The within prepositions, *for*$_{(1.13)}$ and *in*$_{(0.48)}$, satisfy “ntv<jpe.” The within noun, *valuables*$_{(0.47)}$, satisfies “ntv>jpe,” whereas the sentence-final noun, *vault*$_{(1.88)}$, satisfies “ntv<jpe.”

This indicates low power for sentence-final words, whereas high power for sentence-initial words in Japanese English.

### 3.3. Classification of Japanese Subjects’ Proficiencies

For word class, the intensity results in subsection 3.1.1 revealed that Japanese English nouns and /interrogative, negative/ were produced with less intensity, whereas most function words were with more intensity than those by native English.

The power results in subsection 3.1.2 revealed low power for sentence-final words, whereas high power for sentence-initial words in Japanese English.

In this subsection, the Japanese subjects’ English proficiencies are correlated with the results outlined above.

#### 3.3.1. Classification of Japanese subjects

Japanese subjects were classified according to their English proficiency. A native speaker of English, who was an English teacher in Japan, listened to the utterances of each subject, and rated their proficiency. The gauge for determining proficiency level was a ten-point scale, ranging from “fully Japanese spoken English” to “natively spoken English.” Each of the 24 subjects was assigned a level from two to eight. 17 of the subjects were rated at level five and below, and 7 were rated at level six and above. Proficiency at level five and below is referred to as ‘intermediate,’ and proficiency at level six and above is as ‘advanced.’

#### 3.3.2. Word class

Table 7 shows the intensity results, after the subjects’ classification, for nouns, /interrogative, negative/ and function words.

The results suggest that the probability of nouns and /interrogative, negative/ satisfying “ntv>jpe” for the intermediate proficiency level is higher than those for the advanced proficiency level. The results also suggest that the probability of satisfying “ntv<jpe” for function words for the intermediate proficiency level is higher than that for the advanced proficiency level.

These results for intermediate-level and advanced-level Japanese English echo the results for Japanese English and native English in subsection 3.1.1.

Table 8 shows the power results, after the subjects’ classification, for adverbs, /interrogative, negative/ and /conjunction, preposition/.

For word position, the intensity results in subsection 3.2.1 revealed that Japanese English sentence-final words were with less intensity than those by native English.

These results for intermediate-level and advanced-level Japanese English echo the results in subsection 3.1.2.
3.3.3. Word position

Table 9 shows the intensity results, after the subjects’ classification, for word position.

The results suggest that the probability of satisfying “ntv>jpe” for sentence-final words for the intermediate proficiency level is higher than that for the advanced proficiency level.

These data reiterate what was found for Japanese English and native English in subsection 3.2.1.

Table 10 shows the power results, after the subjects’ classification, for word position. The results suggest that the probabilities of satisfying “ntv>jpe” for sentence-initial and within words for the intermediate proficiency level, whereas “ntv<jpe” for sentence-final words for the intermediate proficiency level are higher than those for the advanced proficiency level.

These data reiterate what was found for Japanese English and native English in subsection 3.2.2.

4. DISCUSSION

First, we discuss the results for word class.

The analyses of intensity described in subsection 3.1.1 found that Japanese English nouns and /interrogative, negative/ were produced with less intensity, whereas most function words were with greater intensity than those by native English. As in the analyses after the Japanese subjects’ classification described in subsection 3.2.1, the results correlated with the subjects’ English proficiencies.

Our findings confirmed the results of previous studies [12], that Japanese English content words were uttered non-emphatically. In addition, our findings revealed that Japanese English applies much prominence to function words, which causes the nouns and /interrogative, negative/ to lose their relative prominence.

On another front, the analyses of power in subsection 3.1.2 found that Japanese English adverbs and /interrogative, negative/ were low power, whereas /conjunction, preposition/ were high power. However, in the analyses after Japanese subjects’ classification in subsection 3.3.2, the results did not correlate with the subjects’ English proficiencies.

The differences seen in the results for intensity and power after Japanese subjects’ classification concern the characteristics of Japanese English. The intensity results suggest that English proficiencies are related to power and duration. This supports the findings of previous perception research [13] that indicated that power and duration play a role in English fluency.

Next, we discuss the results for word position.

The analyses of intensity in subsection 3.1.1 found that Japanese English sentence-final words were produced with less intensity than those by native English. In contrast, the analyses of power in subsection 3.2.2 found that Japanese English sentence-initial words were produced with low power, whereas sentence-final words were produced with high power. As in the analyses after Japanese subjects’ classification in subsection 3.3.3, the results correlated with the subjects’ English proficiencies. In addition, our previous study of durations [14] revealed that Japanese English sentence-final words were produced with shorter duration than those by native English.

The results between intensity and power showed differences, whereas the results between intensity and duration showed a similar pattern. This shows that intensity, as defined here, reflects the influence of not power, but duration. In our previous study [14], the short
duration of Japanese English sentence-final words was explained as an irregularity of final lengthening. Similarly, the lower intensity of Japanese English sentence-final words can be explained by the effect of an irregularity of final lengthening on sentence-final word intensity. In contrast, Japanese English initial low power can be explained by the concept of initial strengthening that has been found in English phonetics [15]. That is that Japanese English sentence-initial strengthening does not affect the sentence-initial power peak.

Finally, our findings can be applied to a prosody learning system [16], in which the stressed syllable in a learner’s utterance is detected. Our results indicate that the stressed syllable is a noun, /interrogative, negative/ and the sentence-final word, and that Japanese English does not apply prominence on these words.

The comparison of TIMIT recordings with the English subjects in our study is a first step in establishing equivalency between the two groups. In the next step, we will go into detail. The knowledge gained in our study will be expanded into the relationships between prosodic patterns and human subjects. English subjects for analysis will be grouped into their hometown to gain knowledge on prosodic differences among English dialects. Japanese subjects will be grouped into their language background (e.g., bilingualism or multilingualism) to provide effective knowledge in English learning.

5. CONCLUSION

We have described the characteristics of intensity in Japanese English sentences, relative to word class and word position. Statistical analyses revealed that Japanese English nouns, interrogatives and negatives were produced with less intensity, whereas Japanese English function words were produced with more intensity than those by native English. Another analysis of word position revealed that sentence-final words were produced with less intensity than those by native English. It was also revealed that, for Japanese English, sentence-initial words were produced with low power, whereas sentence-final words were produced with high power.

The knowledge gained from these analyses can be used in English education and computer-based educational systems.

ACKNOWLEDGEMENT

This research is supported in part by a Grand-in-Aid for Research Activity start-up, Project No. 25880025, from the Japan Society for the Promotion of Science.

REFERENCES