# A Study of English Phonetic Characteristics Uttered by Japanese Learners of English in Consideration of Their School Education Background

Tomoko Nariai *

**Abstract**

In an increasingly globalized world, English is growing in importance as an international language, with English education in Japan being reformed to keep up with the trend of allowing students to learn English at a younger age. This study compared the English phonetic characteristics uttered by native speakers of English and Japanese. In the recordings, Japanese subjects said English sentences after listening to the same sentences read by native speakers of English. The results showed that the speech of the Japanese subjects was different from that of the English subjects in all points of the following, intonation, mouth and tongue movement, and duration. The research also suggested that the problem may be caused by a lack of sufficient attention on English pronunciation in schools. Findings indicate that the key for English education reform in Japan is not to reform the guidelines regarding English education in early childhood but to review the recent environment for school education.

**Key words:** speech analysis, fundamental frequency, power spectrum, second language learning, school education

## 1. Introduction

Exposing younger students to more English is seen to help foster internationally competitive talents. Indeed, this belief seems to have driven the Ministry of Education, Culture, Sports, Science and Technology in Japan (MEXT) to make an amendment to the fundamental guidelines of education to let students learn English at a younger age. In the 2020 fiscal year, therefore, English is to become a foreign language class activity for middle elementary grades and will be compulsory for upper elementary grades[1)]. The content of the earlier learning would focus on speaking and listening to English rather than writing and reading English.

GMO Internet Corporation also created a questionnaire survey on attitudes towards English language and distributed it to 10,000 Japanese people[2)]. The results indicated that nine tenths of the Japanese participants consider themselves not to have a good command of English. It can be said that most Japanese people have a feeling of being not proficient in English, which could be a reason

＊　筑波学院大学経営情報学部、Tsukuba Gakuin University

that has driven MEXT to make amendments to the guidelines for education.

The problems of Japanese speakers have been discussed, and this has led to increased knowledge[3–9]. In a study by Kashiwagi and Snyder [3], 40 short sentences uttered by 19 Japanese learners of English were evaluated by three native speakers of English and three native speakers of Chinese. The results suggested that prosody errors of speech did not affect intelligibility for native speakers of English and Chinese. On the other hand, pronunciation errors, especially errors in vowel usage, significantly reduced intelligibility for both groups.

Isida[4] differentiated between two types of artificial English speech: speech with a consonant error and speech with a word accent error. Two speeches were evaluated by 26 native speakers of English and 27 native speakers of Japanese. The speech with the prosodic error influenced intelligibility for the native speakers of English, whereas the speech with the consonant error did not. Conversely, the speech with the consonant error influenced intelligibility for the native speakers of Japanese, whereas the speech with the prosodic error did not.

Sugito[5] suggested that speech with correct intonation sounds more like natural English compared with speech with correct segment.

Derwing, Munro, and Wiebe[6] divided English learners into three groups for a 12-week English program. The first group was educated on pronouncing English segments. The second group was educated on pronouncing English prosody. The third group was educated on the whole of English pronunciation, as is the norm. After the program, a native speaker of English evaluated the proficiency of the learners' English. The results showed that the second group's pronunciation of spontaneous speech improved the most, suggesting that education focused on prosody had a strong influence on pronunciation.

Experts in the field offered a variety of views. Roach[7] suggested that Japanese speakers of English were not good at schwas and vowels, replacing every schwa with a Japanese vowel in their utterances. That turned the English voiced by Japanese speakers into "katakana English," which sounds unnatural to native speakers of English. Yabuuchi and Satoi[8] suggested that the Japanese would sound more natural when speaking English if they spoke it with a dynamic pitch range. Yamane[9] suggested that English speech in which Japanese speakers made word accent errors reduced intelligibility by native speakers of English to a level of 47.5%. Much knowledge has been garnered regarding the characteristics of Japanese speakers of English, and three key problems have consequently been highlighted.

One problem is that most of the studies only investigated English proficiency, meaning that phonetic realizations have not yet been examined sufficiently. Another problem is that most studies aimed to pass on the knowledge they obtained on English education, yet they only investigated the English skills of Japanese subjects by referring to the grade of English proficiency tests or scores on the Test of English for International Communication. This approach failed to take into account that a person's English skills depend on what kind of education they receive. The final problem is that most studies did not make it clear that the irregularities in English pronunciation occurred because the subjects did not know how to pronounce English or could not pronounce English.

This study investigated the phonetic characteristics uttered by Japanese learners of English,

considering the relation between the subject's educational background and their English proficiency. In order to clarify why the irregularities in English pronunciation occur, subjects were presented with the relevant English sentences prior to recording and were asked to practice uttering them. In the recording, Japanese subjects read the English sentences orally after listening to the same sentences read by native speakers of English.

Section 2 describes the material used in the study, while section 3 shows the results. Discussion on relevant issues and conclusions comprise sections 4 and 5, respectively.

## 2. Materials

### 2.1. Speech materials

The subjects consisted of six male participants. Three were native speakers of Japanese, aged 20 to 21. There were six sentences that were chosen from MOCHA-TIMIT[10]. Those were timit070, 200, 221, 228, 268, 392, and 460. The readings of the Japanese subjects were recorded on Mondays in November, 2018.

Another three subjects were native speakers of English whose speeches were collected in a previous research project[11]. These were labeled with the sentence numbers English subject1, English subject2, and English subject3, respectively.

### 2.2. Recording conditions

The Japanese subjects were asked to practice reading the English sentences prior to recording. The recording was done in the computer exercise room at Tsukuba Gakuin University. The recording device was a Sony linear PCM recorder, PCM-D100. The speeches were recorded through 2 channels at sampling of 16 bit and 48,000 Hz.

The recording procedures were as follows.

1. the subject sat in a designated seat
2. instructions were provided regarding the recording devices
3. the subject was instructed to maintain distance between his mouth and the recorder on the table
4. the subject was instructed on how to listen to the sentences read by the native speakers of English
5. finally, the subject was instructed on how to record using the given recording devices.

Subjects were required to utter each sentence repeatedly until the speech sample was recorded properly.

In the order they finished, speeches were labeled with the sentence numbers Japanese subject1, Japanese subject2, Japanese subject3, respectively.

For analysis, the speeches were edited using the speech editing software, *Audacity* [12], in order to cut the abundant sound and change 1 channel.

### 2.3. Speech analysis

Using a tool for sound visualization and manipulation, *Wavesurfer*[13], each speech was analyzed to extract the spectrogram, with the first three vowel formant track superimposed with a window length of 25 ms and a frame shift of 5 ms.

For the three formants, the fundamental frequency (F0, henceforth) is related to vowel height. F1 is related to the degree of backness of tongue, where, the higher the F1, the more frontal the vowel. F2 is related to the degree of lip-rounding, where, the lower the F2, the rounder the shape of the lip. The formant sequence was then manually segmented into a word sequence.

The power pattern was also 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 power value was 80 dB, and the minimum was 0 dB. The power value sequence was also manually segmented into a word sequence.

For each word of the speech, the following six parameters were extracted.

・F0 peak (Hz) : the peak value of F0 for the word
・Power peak (dB) : the peak value of power for the word
・F1 range (Hz) : the peak F1 value minus the lowest F1 value for the word
・F2 range (Hz) : the peak F2 value minus the lowest F2 value for the word
・Duration (sec) : the duration of the word

## 3. Result

In this section, the speech of the English subjects and the Japanese subjects was compared. In advance, the Japanese were asked to answer questions about their English learning backgrounds.

### 3.1. Results from the interview with Japanese subjects

Research was undertaken using a questionnaire survey on the English learning background of each Japanese subject. The four questions that the subjects were asked to answer are listed below. The Common European Framework of Reference for Languages (CEFR, henceforth)[14] was used to assess English skill by self-reported number. The CEFR English language level description ranges from A1, beginner, to C2, proficient English user. The handouts of CEFR were distributed to the subjects.

**Q1.** What is your CEFR English level[14] ?
**Q2.** How long have you been studying English at school or college?
**Q3.** Have you been taught how to intonate English when you pronounce a sentence in school?
**Q4.** Have you been taught how to move your mouth and tongue to pronounce a word in school?

The answers to the questions are shown in **Table** 1.

Table 1. Results of the questions asked to the Japanese subjects

| | Q1 | Q2 | Q3 | Q4 |
|---|---|---|---|---|
| **Japanese subject1** | A1 | 11 years | No | No |
| **Japanese subject2** | A1 | 11 years | No | No |
| **Japanese subject3** | A1 | 9 years | No | No |

### 3.2. Results of speech analysis

**Table 2-1** shows the results of English subject1 in sentence 026, "Most young rabbits rise early every morning." The column indicates each word. The row indicates the F0 peak for word, power peak for word, F1 range, F2 range, and duration, respectively.

In English subject1, the F0 peak in "most" continuously declines to "rise" and then rises in "early," then continuously declines to "morning." The power peak in "most" continuously rises in "rabbits" and then continuously declines to "early," then rises to "every morning."

**Table 2-2** shows the results of Japanese subject1 in sentence 026. The F0 peak in "most" continuously declines to "rabbits" and then rises in "rise," then continuously declines to "every," then rises in "morning." The power peak in "most" continuously declines to "young" and rises in "rabbits," then continuously declines to "morning."

In **Table 2-1**, the F1 range is large in "most" and "rabbits" but short in "early" and "every." The F2 range is large in "most," "rabbits," and "morning" but short in "rise" and "early." The duration is long in "rabbits" and "every" but short in "young" and "early."

In **Table 2-2**, the F1 range is large in "early" and "morning" but short in "young" and "every." The F2 range is large in "early," "every," and "morning" but short in "young" and "rise." The duration is long in "rise" and "rabbits" but short in "young" and "most."

Table 2-1. Results of English subject1 in TIMIT 026

| | Most | young | rabbits | rise | early | every | morning |
|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 168.76 | 124.62 | 105.58 | 99.19 | 108.28 | 98.36 | 93.82 |
| **Power peak(dB)** | 60.63 | 62.14 | 63.30 | 62.85 | 58.77 | 59.18 | 59.18 |
| **F1 range(Hz)** | 1158.34 | 333.77 | 1119.50 | 394.66 | 117.10 | 262.24 | 324.09 |
| **F2 range(Hz)** | 1775.53 | 876.83 | 1429.23 | 575.13 | 663.10 | 1170.10 | 1520.67 |
| **Duration(sec)** | 0.40 | 0.20 | 0.53 | 0.33 | 0.32 | 0.41 | 0.36 |

Table 2-2. Results of Japanese subject1 in TIMIT 026

| | Most | young | rabbits | rise | early | every | morning |
|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 178.00 | 149.00 | 126.00 | 138.00 | 130.00 | 118.00 | 130.00 |
| **Power peak(dB)** | 60.32 | 57.56 | 58.99 | 58.64 | 54.07 | 54.07 | 51.35 |
| **F1 range(Hz)** | 750.03 | 351.50 | 525.09 | 809.05 | 1373.26 | 446.84 | 1231.60 |
| **F2 range(Hz)** | 1526.12 | 481.21 | 1265.30 | 956.08 | 2174.60 | 1896.97 | 1747.00 |
| **Duration(sec)** | 0.35 | 0.33 | 0.63 | 0.67 | 0.38 | 0.37 | 0.56 |

**Figure 1-1** and **Figure 1-2** show the results of English subject1 and Japanese subject1,

respectively. The *x*-axis shows each word for sentence 026. The *y*-axis 1 indicates the F0 peak and power peak, and the *y*-axis 2 indicates F1 range and F2 range.

In F0 peak, the figures show that English subject1 and Japanese subject1 differ in a verb "rise." English subject1 declines, while Japanese subject1 rises.

There are also differences in F1 range and F2 range between English subject1 and Japanese subject1. English subject1's mouth and tongue moved in "most" and "rabbits" but did not move in "early." Meanwhile, Japanese subject1's mouth and tongue moved in "early" and "morning."

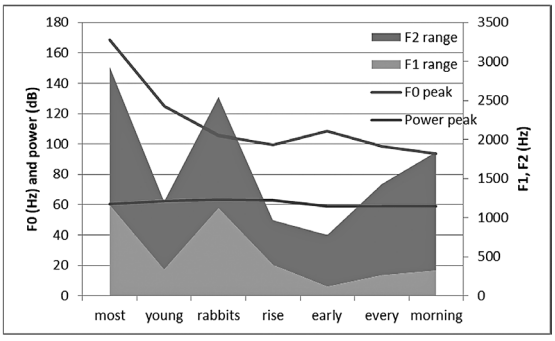


Figure 1-1. Results of English subject1 in TIMIT 026

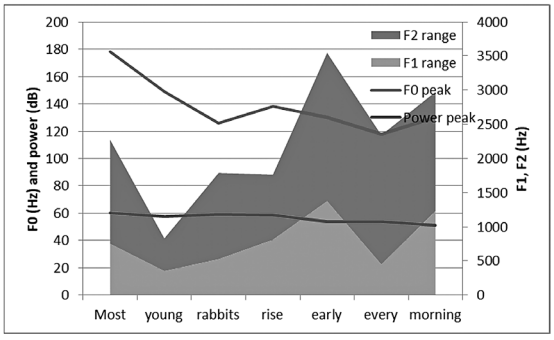


Figure 1-2 Results of Japanese subject1 in TIMIT 026

**Table** 3-1 and **Table** 3-2 show the results of English subject1 and Japanese subject1 in sentence 109, "Birthday parties have cup cakes and ice cream."

Table 3-1. Results of English subject1 in TIMIT 109

| | Birthday | parties | have | cup | cakes | and | ice | cream |
|---|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 147.83 | 111.1 | 106.0 | 153.9 | 146.1 | 110.62 | 120.3 | 110.08 |
| **Power peak(dB)** | 58.80 | 54.82 | 52.18 | 57.50 | 55.26 | 48.76 | 62.31 | 43.86 |
| **F1 range(Hz)** | 370.81 | 420.7 | 720.9 | 789.1 | 1123.6 | 378.11 | 559.4 | 1299.9 |
| **F2 range(Hz)** | 997.32 | 965.8 | 497.2 | 943.2 | 949.3 | 652.5 | 466.6 | 1036.8 |
| **Duration(sec)** | 0.58 | 0.41 | 0.19 | 0.2 | 0.44 | 0.17 | 0.27 | 0.33 |

Table 3-2. Results of Japanese subject1 in TIMIT 109

| | Birthday | parties | have | cup | cakes | and | ice | cream |
|---|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 175.64 | 160.7 | 110.21 | 108.5 | 163.5 | 109 | 118.9 | 143.9 |
| **Power peak(dB)** | 55.50 | 57.68 | 49.27 | 62.6 | 56.17 | 52.6 | 61.40 | 46.76 |
| **F1 range(Hz)** | 490.49 | 461.8 | 634.7 | 810.6 | 270.0 | 530 | 1100.9 | 777.7 |
| **F2 range(Hz)** | 1266.32 | 595.1 | 427.9 | 1387 | 331.9 | 707 | 1597 | 949.1 |
| **Duration(sec)** | 0.56 | 0.49 | 0.35 | 0.25 | 0.26 | 0.32 | 0.37 | 0.32 |

**Figure 2-1** and **Figure 2-2** show the results of English subject1 and Japanese subject1, respectively.

In terms of F0 peak, English subject1 and Japanese subject1 differ in "cup," "cakes," "ice," and "cream." In "cup," English subject1 rises, while Japanese subject1 declines. In "cakes," English subject1 declines, while Japanese subject1 rises. In "ice," English subject1 rises, while Japanese subject1 declines. In "cream," English subject1 declines, while Japanese subject1 rises.

In the F1 range and the F2 range, English subject1's mouth and tongue moved in "cakes" and "cream." By contrast, Japanese subject1's mouth and tongue moved in "cup" and "ice."

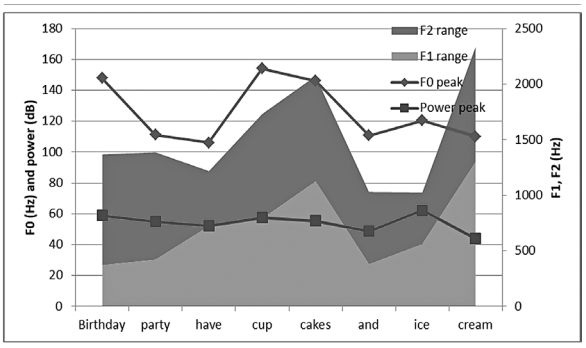


Figure 2-1. Results of English subject1 in TIMIT 109

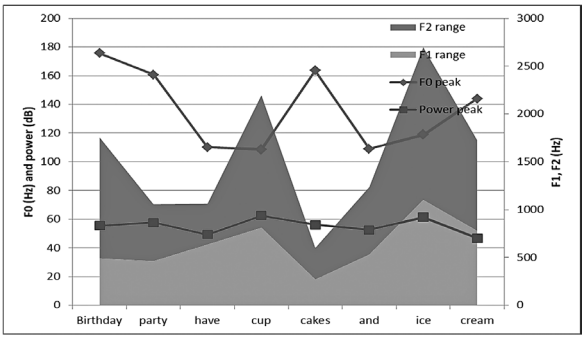


Figure 2-2. Results of Japanese subject1 in TIMIT 109

**Table** 4-1 and **Table** 4-2 show the results of English subject2 and Japanese subject2 in sentence 119, "The mango and the papaya are in a bowl."

Table 4-1. Results of English subject2 in TIMIT 119

| | **The** | **mango** | **and** | **the** | **papaya** | **are** | **in** | **a** | **bowl** |
|---|---|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 112 | 163 | 109 | 106 | 160 | 107 | 109 | 103 | 146 |
| **Power peak(dB)** | 50 | 63 | 53 | 47 | 62 | 51 | 50 | 47 | 54 |
| **F1 range(Hz)** | 282 | 440 | 412 | 567 | 1027 | 268 | 141 | 341 | 874 |
| **F2 range(Hz)** | 475 | 1316 | 254 | 573 | 1756 | 653 | 897 | 263 | 1691 |
| **Duration(sec)** | 0.11 | 0.53 | 0.13 | 0.17 | 0.52 | 0.15 | 0.14 | 0.6 | 0.48 |

Table 4-2. Results of Japanese subject2 in TIMIT 119

| | **The** | **mango** | **and** | **the** | **papaya** | **are** | **in** | **a** | **bowl** |
|---|---|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 150 | 185 | 139 | 141 | 153 | 134 | 138 | 124 | 181 |
| **Power peak(dB)** | 53 | 51 | 51 | 48 | 52 | 52 | 45 | 50 | 46 |
| **F1 range(Hz)** | 1314 | 785 | 1073 | 975 | 717 | 346 | 346 | 154 | 291 |
| **F2 range(Hz)** | 1248 | 668 | 1231 | 764 | 1259 | 341 | 1501 | 325 | 865 |
| **Duration(sec)** | 0.42 | 0.7 | 0.45 | 0.41 | 0.63 | 0.37 | 0.3 | 0.2 | 0.44 |

**Figure** 3-1 and **Figure** 3-2 show the results of English subject2 and Japanese subject2.

In terms of F0 peak and power peak, the figures differ in nouns. English subject2 rises in "mango," "papaya," and "bowl," equally. For Japanese subject2, F0 peak rises in "mango" and "bowl" but declines in "papaya," and power peak falls on "bowl."

In the F1 range and the F2 range, English subject2's mouth and tongue moved in "mango," "papaya," and "bowl," while Japanese subject2's mouth and tongue moved in "The" and "and."

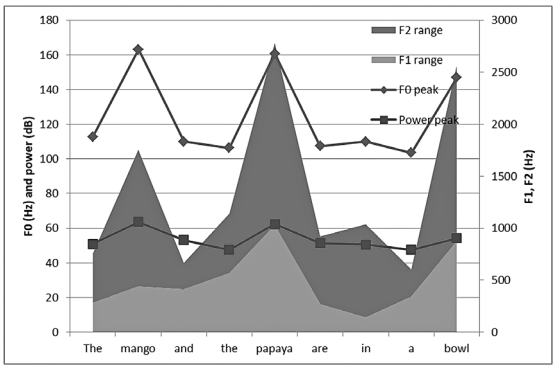


Figure 3-1. Results of English subject2 in TIMIT 119

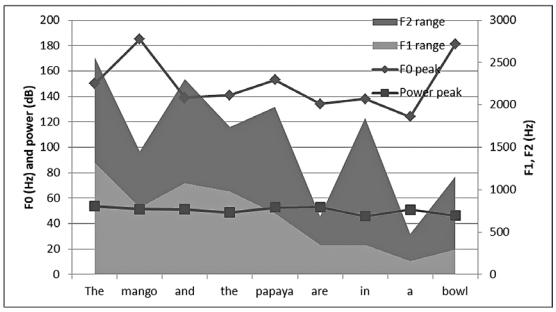


Figure 3-2. Results of Japanese subject2 in TIMIT 119

**Table 5-1** and **Table 5-2** show the results of English subject2 and Japanese subject2 in sentence 268, "How oily do you like your salad dressing?"

Table 5-1. Results of English subject2 in TIMIT 268

| | **How** | **oily** | **do** | **you** | **like** | **your** | **salad** | **dressing** |
|---|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 186 | 204 | 132 | 135 | 127 | 123 | 136 | 106 |
| **Power peak(dB)** | 63 | 59 | 60 | 55 | 60 | 66 | 66 | 65 |
| **F1 range(Hz)** | 625 | 266 | 114 | 61 | 420 | 218 | 796 | 477 |
| **F2 range(Hz)** | 978 | 1535 | 284 | 690 | 1209 | 285 | 755 | 527 |
| **Duration(sec)** | 0.19 | 0.35 | 0.09 | 0.07 | 0.29 | 0.1 | 0.33 | 0.42 |

Table 5-2. Results of Japanese subject2 in TIMIT 268

| | How | oily | do | you | like | your | salad | dressing |
|---|---|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 164 | 195 | 146 | 152 | 140 | 137 | 155 | 188 |
| **Power peak(dB)** | 55 | 54 | 48 | 49 | 53 | 54 | 58 | 61 |
| **F1 range(Hz)** | 693 | 212 | 417 | 114 | 395 | 561 | 876 | 882 |
| **F2 range(Hz)** | 1007 | 1513 | 727 | 618 | 772 | 1178 | 485 | 1182 |
| **Duration(sec)** | 0.56 | 0.73 | 0.3 | 0.25 | 0.36 | 0.49 | 0.57 | 0.65 |

**Figure** 4-1 and **Figure** 4-2 show the results of English subject2 and Japanese subject2.

In terms of F0 peak, the figures differ in "salad" and "dressing." English subject2 rises in noun, "salad," and declines in "dressing," while Japanese subject2 continuously rises in "salad dressing."

In the F1 range and the F2 range, English subject2's mouth moved in "How oily," "like," and "salad," while Japanese subject2's mouth moved in "How oily," "your," and "dressing."

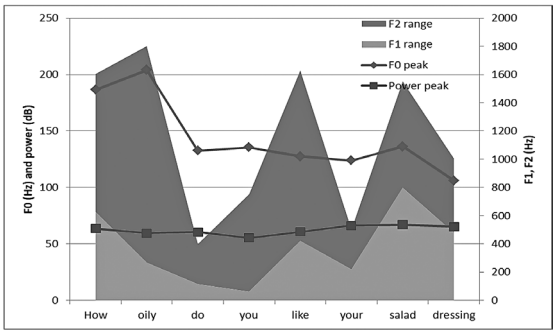


Figure 4-1. Results of English subject2 in TIMIT 268

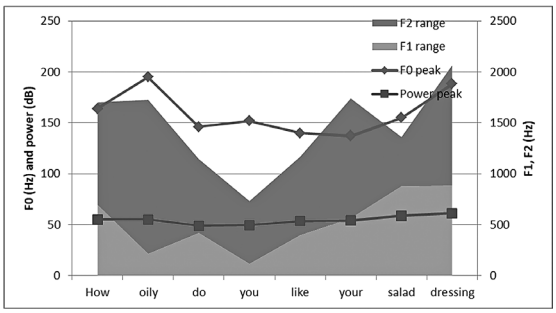


Figure 4-2. F0 and power of English group in sentence of 268s

**Table** 6-1 and **Table** 6-2 show the results of English subject3 and Japanese subject3 in sentence 215, "Only the best players enjoy popularity."

Table 6-1. Results of English subject3 in TIMIT 215

| | Only | the | best | players | enjoy | popularity |
|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 157.61 | 106.09 | 130.72 | 134.27 | 94.40 | 100.40 |
| **Power peak(dB)** | 68.68 | 52.11 | 64.45 | 65.56 | 58.54 | 57.68 |
| **F1 range(Hz)** | 1066.85 | 192.74 | 1406.92 | 782.44 | 337.69 | 505.49 |
| **F2 range(Hz)** | 1337.69 | 1402.24 | 974.92 | 1255.47 | 1237.00 | 1101.58 |
| **Duration(sec)** | 0.24 | 0.13 | 0.23 | 0.44 | 0.32 | 0.7 |

Table 6-2. Results of Japanese subject3 in TIMIT 215

| | Only | the | best | players | enjoy | popularity |
|---|---|---|---|---|---|---|
| **F0 peak(Hz)** | 157.61 | 134.27 | 106.16 | 100.40 | 125.00 | 166.00 |
| **Power peak(dB)** | 68.68 | 64.45 | 65.56 | 57.68 | 50.46 | 44.01 |
| **F1 range(Hz)** | 1349.42 | 1406.92 | 387.46 | 1386.61 | 165.00 | 952.71 |
| **F2 range(Hz)** | 1431.52 | 1501.81 | 1237.00 | 1175.19 | 484.95 | 1598.12 |
| **Duration(sec)** | 0.55 | 0.41 | 0.64 | 0.79 | 0.27 | 1.01 |

**Figure 5-1** and **Figure 5-2** show the results of English subject3 and Japanese subject3.

In terms of F0 peak, the figures differ in "the best players" and "enjoy popularity." English subject3 continuously rises in noun, "the best players," then declines in "enjoy" and rises in "popularity," while Japanese subject3 continuously declines in "the best players" and continuously rises in "enjoy popularity."

In the F1 range and the F2 range, English subject3's mouth moved in "only" and "best," while Japanese subject3's mouth moved in "The" and "players."

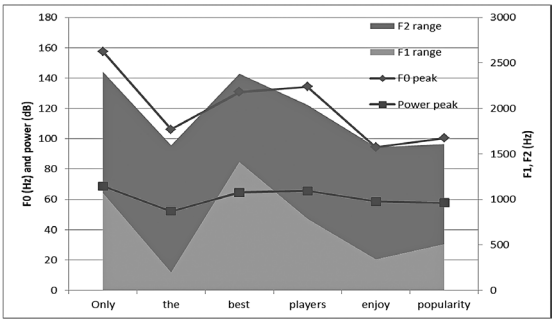


Figure 5-2 Results of English subject3 in TIMIT 215

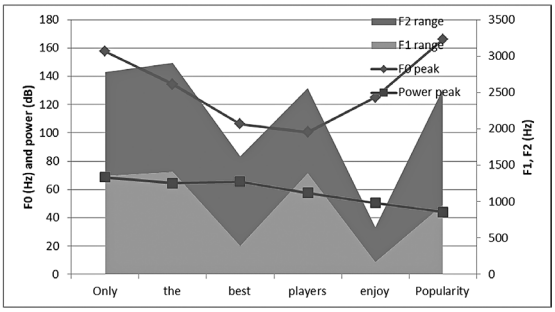


Figure 5-2 Results of Japanese subject3 in TIMIT 215

**Table 7-1** and **Table 7-2** show the results of English subject3 and Japanese subject3 in sentence 221, "How permanent are their records?"

Table 7-1. Results of English subject3 in TIMIT 221

| | **How** | **permanent** | **are** | **their** | **records** |
|---|---|---|---|---|---|
| **F0 peak(Hz)** | 161.80 | 164.86 | 114.10 | 112.86 | 118.95 |
| **Power peak(dB)** | 63.38 | 59.79 | 63.40 | 56.71 | 62.44 |
| **F1 range(Hz)** | 864.81 | 1031.84 | 421.94 | 276.30 | 1281.59 |
| **F2 range(Hz)** | 807.66 | 855.03 | 445.47 | 608.96 | 1627.44 |
| **Duration(sec)** | 0.27 | 0.44 | 0.26 | 0.21 | 0.65 |

Table 7-2. Results of Japanese subject3 in TIMIT 221

| | **How** | **permanent** | **are** | **their** | **records** |
|---|---|---|---|---|---|
| **F0 peak(Hz)** | 155.00 | 140.00 | 124.00 | 124.00 | 134.00 |
| **Power peak(dB)** | 54.64 | 63.04 | 50.43 | 53.12 | 54.78 |
| **F1 range(Hz)** | 567.59 | 893.76 | 573.19 | 461.24 | 777.69 |
| **F2 range(Hz)** | 493.95 | 1288.78 | 807.79 | 657.38 | 1022.18 |
| **Duration(sec)** | 0.25 | 0.63 | 0.23 | 0.31 | 0.75 |

**Figure 6-1** and **Figure 6-2** show the results of English subject3 and Japanese subject3.

In terms of F0 peak, the figures differ in "permanent." English subject3 continuously rises in "How permanent" and declines in "are," while Japanese subject3 continuously declines in "How permanent are."

In the F1 range and the F2 range, English subject3's mouth and tongue moved in "record," while Japanese subject3's mouth and tongue moved in "permanent."

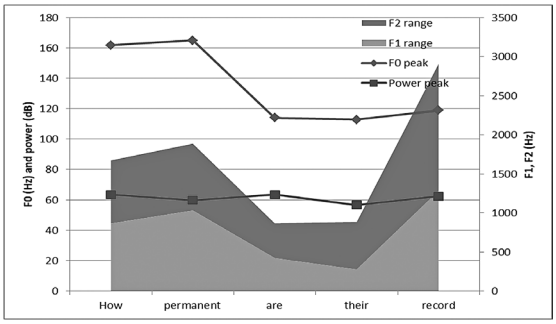


Figure 6-1 Results of English subject3 in TIMIT 221

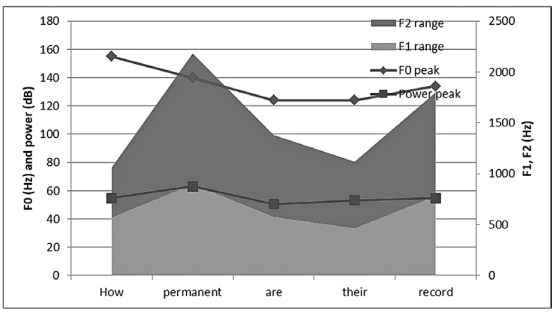


Figure 6-2 Results of Japanese subject3 in TIMIT 221

### 3.3. Discussions

In subsection 2.1, the English education background of the Japanese subjects was investigated. The subjects' English language level was beginner according to self-assessment. Their periods of English learning ranged from 9 to 11 years, which suggested that all the subjects had studied English at elementary school or junior high school in Japan. They had not, however, been educated on how to pronounce and intonate English in detail during their school days.

In subsection 2.2, the characteristics of English subjects and Japanese subjects were compared. The two groups differ in the following two points.

First, English subjects declined F0 peaks in a verb, auxiliary verb, or postpositional particle, while Japanese subjects did not. For sentence 026, English subject phrased at "Most young rabbits rise" and "early every morning," while Japanese subject did not. For sentence 109, English subject phrased at "Birthday parties have," "cup cakes and," and "ice cream," while Japanese subject did not. For sentence 119, English subject declined at "The mango and the," "papaya are in a," and "bowl." For sentence 268, English subject phrased at "How oily do," "you like your," and "salad dressing," while Japanese subject did not. For sentence 215, English subject phrased at "Only the," "best players enjoy," and "popularity," while Japanese subject did not. For sentence 221, English subject phrased in "How permanent are their" and "record," while Japanese subject did not.

Second, English subjects moved their mouth and tongue for some words, while Japanese subjects did so for other words. Mouth and tongue movement is complicated and related to the other parameters. Such details will consequently be discussed in future research.

## 4. Discussions

In subsection 3.2, the English of the Japanese subjects was compared to that of the English subjects. The results suggest that the Japanese subjects differed from the English subjects in all points: intonation, mouth and tongue movement, and duration.

In subsection 3.1, research using a questionnaire survey on the English learning backgrounds of the Japanese subjects was presented. All the Japanese subjects had studied English at elementary or junior high school. They had not, however, been educated on how to pronounce and intonate English in detail in their school days. This knowledge confirms Tejima's comment [15] that, in English education at junior high school, vowel and consonant pronunciation guidance has rarely been provided.

Such condition do not appear to uphold the guidelines regarding English education in junior high schools and high schools issued by MEXT [16, 17]. In the revision made in 1998, at which time the Japanese subjects were of school age, the aims of English education were for students to become accustomed to the basic characteristics of English, and to say English words and sentences properly. This included pronunciation, linking, accent, intonation, and pauses. In the education guidelines for high schools, the aims of English education have been expanded to include the ability to pronounce English in detail with appropriate rhythm, intonation, loudness, and speed. This approach therefore suggests that the educational aims for English education determined by MEXT were intended to cover the whole spectrum of English pronunciation.

The problem for Japanese people with poor English skills becomes clear: It is based on school education. The guidelines for English education in high schools as determined by MEXT cover English pronunciation in its entirety. The Japanese subjects were different from the English subjects in all points concerning pronunciation because they had never been taught how to speak English in school and did not know how to pronounce the language.

MEXT seems determined to amend the fundamental law of education to allow students to learn English at a younger age [1]. In English education, the quality - the content of the class - must be more important than the quantity - the length of the class. The key for English education reform would not, therefore, be reforming the guidelines for English education in early childhood but rather reviewing the recent environment for foreign language education of the junior and senior high school levels.

## 5. Conclusion

This study investigated the English phonetic characteristics uttered by Japanese subjects in consideration of their English education background. Before recording, the Japanese subjects

practiced the English sentences after listening to the speeches uttered by native speakers of English. The speech of the Japanese subjects was, however, different from that of the English subjects in all points. The Japanese subjects had not been taught how to intonate English or move their mouth and tongue in English in their school days. Their performance was therefore caused by a lack of sufficient education in English pronunciation according to the guidelines in their school days.

## 6. Reference


1) MEXT, Foreign Language Education [http://www.mext.go.jp/a_menu/kokusai/gaikokugo/index.htm] Oct4,2019.
2) GOM, 英語に関する意識調査 [https://prtimes.jp/main/html/rd/p/000002228.000000136.html] Dec 3, 2018.
3) A. Kashiwagi, M. Snyder, "Intelligibility of Japanese College Freshmen as Listened to by Native and Nonnative Listeners," Jacet Journal, Vol.53, pp39–56, 2014.
4) K. Ishida, "Factors affecting intelligibility of English: comparison of effects caused by segmental and prosody errors generated by the text-to-speech technology," Journal of Kansai University Graduate School of Foreign Language Education and Research, Vol.11, pp71–97, 2013.
5) M.Sugito, Nihon jin no Eigo,Izumi Shoin, 1996.
6) T. Derwing, M. Munro, G. Wiebe, "Evidence in favor of a broad framework for pronunciation instruction", Language Learning, Vol.48, pp393–410,1998.
7) P. Roach, "English Phonetics and Phonology: A Practical Course," Cambridge University Press, 2009.
8) S. Yabuuchi, H. Satoi, "Prosodic Characteristics of Japanese EFL Learners' Oral Reading Comparison between Good and Poor Readers,"Gaikokugo Kyoiku Media gakkai kikan shi, Vol.38, pp99–112, 2001.
9) S. Yamane, Notes on the intelligibility of Japanese EFL learners' pronunciation, Kansai Daigaku Gaikokugo Gakubu Kiyo, Vol. 13, pp129–141, 2015.
10) University of Edinburgh, Mocha-Timit [http://www.cstr.ed.ac.uk/research/projects/artic/mocha.html], Oct. 4, 2019.
11) K. Tanaka, Y. Itoh, H. Kojima, N. Fujimura, "Speech data retrieval system constructed on a universal phonetic code domain," ASRU 2001, pp. 323–326, 2001.
12) Audacity, [https://www.audacityteam.org/], Oct 4, 2019.
13) Royal Institute of Technology, Wavesurfer, [https://www.speech.kth.se/wavesurfer/], Oct 4, 2019.
14) Cambridge Centre, CEFR [https://cambridgecentre.jp/exams/cefr/], Oct 4, 2019.
15) M.Tejima, "On Teaching English Pronunciation in Japanese Secondary Education How It Is and How It Should Be," Onsei Kenkyu, Vol.15, No.1, pp31–43, 2011.
16) MEXT, Chapter 2 subject, Section 9 foreign language [http://www.mext.go.jp/a_menu/shotou/cs/1320086.htm]
17) MEXT, Chapter 2 subject for education, Section 8, foreign language, [http://www.mext.go.jp/a_menu/shotou/cs/1320334.htm]