Correlation between schwa pronunciation and perceived speaking skill of Japanese EFL learners

Miyuki Sakuma s1160091

Supervised by Professor Ian Wilson

Abstract

Schwa sound is difficult to pronounce for Japanese because there is no such vowel sound in Japanese. In this paper, I investigated whether studying schwa is important or not and found out which properties of schwa most affect the perception of Japanese EFL learners’ English pronunciation. This research will help Japanese EFL learners to improve their English pronunciation. I analyzed the learners’ schwa sounds by using Praat software. I found there is a strong correlation between perceived English pronunciation skill and duration of schwa, namely, the shorter the duration of schwa, the higher their perceived pronunciation level.

1 Introduction

Most EFL learners have trouble in English pronunciation, because the sound system of Japanese is very different from that of English. Japanese moras consist of either a vowel or set of a consonant and a vowel, except for geminate consonant /Q/ and moraic nasal /N/. Furthermore, there is pitch accent in Japanese but there is no stress accent, and English vowels are difficult to pronounce for Japanese EFL learners. These features have affected Japanese learners’ pronunciation.

In past research, Satoi [?] compared the weak vowels spoken by Japanese EFL learners with those spoken by native English speakers. Japanese EFL learners pronounce schwa that has a greater distance from the centroid in the F1 and F2 formant space than does that of native English speaker. In addition, Sugai [?] found that the Japanese learners’ vowel pronunciation was strongly influenced by the spelling. This means that Japanese learners often pronounce schwa like the ROMAJI reading for the vowel that should have been reduced to schwa.

Schwa sound is difficult to pronounce correctly for Japanese because there is no such sound in Japanese. Schwa sound is a weak vowel, appears in unstressed syllables, and is characterized by these properties: shorter in duration, quieter (lower intensity), and lower in pitch. Figure 1 shows the phonetic symbols of American English vowels and where they are located in the vowel space. Note that schwa is in the center of the vowel space. It is regrettable that there are few Japanese EFL learners who know schwa sound because they do not learn about this sound in school. The purpose of this thesis research is to investigate whether studying schwa is important or not for the learners, and to determine which properties most strongly correlate with good perceived pronunciation. We checked the pronunciation skills of Japanese learners of English, especially regarding schwa. The properties we investigated are duration, pitch, intensity, F1 and F2. Moreover, the learners were judged on their level of English pronunciation by native listeners and I researched whether there is a correlation between the properties and the perceived pronunciation level. This enabled us to figure out what the biggest problem is of schwa pronunciation for Japanese EFL learners. I hope this research will help Japanese learners who want to improve their proficiency in English pronunciation.

Figure 1: Phonetic symbols of American English vowels and their position in the vowel space
2 Method

2.1 Subjects

In this experiment, sound data were collected from 32 native Japanese men, but 3 subjects’ data were not used because the audio quality was not good enough. In addition, there was an extreme outlier who was found when making graphs, and so that speaker was excluded. His pitch and intensity data was left out of the group.

Subjects did not know the purpose of this experiment. The learners read an English paragraph aloud. The paragraph of this survey is shown in the Appendix. Then their English pronunciation levels were judged by 4 native North American English listeners. The judges were not told what aspect of pronunciation to focus on, so they were probably not focused only on schwa.

2.2 Data collection

I used the paragraph ‘The North wind and the Sun’. It was proved useful in illustrating the characteristics of basic speech sounds in different languages [7]. The recording was conducted in a classroom using a headset microphone (Sennheiser PC 131), an iMac computer (Mac OS 10.6.8), and Praat (version 5.2.38). First, the learners silently read the paragraph that was printed on paper. Then they read the paragraph aloud and recorded it. There were three subjects who had unclear data and it was thought that their microphone settings had been incorrect.

Judges listened to a sound file which consisted of each subject’s first sentence. This sound file was prepared using Praat. There was a 4-second interval between each sentence and judges judged the subjects’ English pronunciation level at this time. The level was judged on a Likert scale from 1 (very poor) to 7 (native) and the judges chose the level by circling a number on a mark sheet.

2.3 Apparatus

Praat was used to record the voice and to edit the file. Sound files were saved as WAV files.

Microsoft Word was used to make the English level judgement form.

Microsoft Excel was used to organize data, do statistics, and make graphs.

2.3.1 Data Analysis

I analyzed the sound data using Praat. I opened the sound data files with Praat and displayed a spectrogram. Then I annotated the part of schwa by using TextGrid like Figure 2. When choosing the part of schwa to label, I observed how formants were changing. Figure 2 shows the word “disputing”, and you can see the difference of the formants between schwa and the sounds around it. I analyzed 10 occurrences of schwa per subject, those shown in Table 1. I measured the duration, pitch, intensity, F1 and F2 of each schwa and recorded the data in an Excel file.

<table>
<thead>
<tr>
<th>Spell</th>
<th>Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>along, agreed, around, attempt, immediately</td>
</tr>
<tr>
<td>i</td>
<td>disputing</td>
</tr>
<tr>
<td>u</td>
<td>succeeded</td>
</tr>
<tr>
<td>e</td>
<td>traveller</td>
</tr>
<tr>
<td>o</td>
<td>considered, confess,</td>
</tr>
</tbody>
</table>

Table 1: Schwa

Measuring schwa pitch had a problem that there were some undefined pitch values. It was because schwa sound is so short. In those cases, I removed those schwas when calculating the mean pitch of schwa. There are no more than 1or 2 undefined schwas per learner, and so these did not have a large effect.

The formula (7) was used to estimate the distance of schwa vowel to the mid-center of the subject’s vowel space. $D_{k}$ is the dispersion the $k^{th}$ token in vowel space $v$. $v_{1k}$ is the F1 value for the $k^{th}$ token.
And I calculated MeanF. First, I annotated /a/, /i/, and /u/ by using a TextGrid, and measured F1 F2 at the center of the vowel duration. Table 2 shows the stimuli that was used when calculating MeanF. When annotating /i/, there was a problem that some learners read ‘succeeded’ incorrectly as ‘success’. Then, ‘agreed’ was used in those cases.

$$D_{ik} = \sqrt{(MeanF1 - v_{1k})^2 + (MeanF2 - v_{2k})^2} \quad (1)$$

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>along(o)</td>
</tr>
<tr>
<td>/i/</td>
<td>succeeded(e)</td>
</tr>
<tr>
<td>/u/</td>
<td>two(wo)</td>
</tr>
</tbody>
</table>

Table 2: Mean F1 F2

3 Results

Four types of correlation graphs are shown as follows. Figure 3 shows the correlation between perceived pronunciation skill and percentage of schwa in the length of whole sentence. It can be seen that duration of schwa is negatively correlated with English level ($r = -0.54$).

Using Cohen’s [?] guideline for reporting behavioural science effect sizes (small effect size, $r = 0.1 - 0.23$; medium, $r = 0.24 - 0.36$; large, $r = 0.37$ or larger), duration of schwa have large effect on perceived speaking skill. There is small correlation with perceived pronunciation skill and ration of schwa pitch to mean pitch of paragraph in Figure 4. In figures 5 and 6, there is no correlation.

4 Discussion

In the experiment, it was confirmed that schwa is important for perceived pronunciation skill. The duration of schwa is strongly correlated with perceived pronunciation level of English. This result is a clue how to study schwa for Japanese EFL learner.

There are some values that are higher than 1 on the horizontal axis in Figure 4. Such values mean the learner’s schwa pitch is the same or higher than the whole paragraph. Furthermore, there are a few scattered values further to the right. What these values mean is that the learner put the accent on the part of schwa.
In Figure 5, the values cluster together and are higher than 1 of the horizontal axis. This is interesting data in that all learners pronounce schwa intensity louder than the average intensity of the whole paragraph. Altogether then, the learners pronounced schwa with a higher pitch and louder than a native English speaker would.

Figure 6 shows correlation between pronunciation skill and distance of schwa to mid-center of the vowel space. I expected that there would be a correlation. However, there is no correlation and the values disperse widely. It is possible that schwa was too short for native listeners to get much information about the formant values, and so native listeners did not make their pronunciation judgements based on vowel space location.

5 Conclusions and Future Work

The results obtained by this experiment can help Japanese EFL learners who are troubled by English pronunciation. Focusing on making the length of schwa very short is expected to improve Japanese EFL learners’ English pronunciation skill.

Moreover, other problems with the learner’s schwas were found. As explained in the introduction to this paper, schwa sound has a low pitch. However, most learners pronounced schwa the same pitch as the mean pitch of the paragraph. In addition, their English was loud intensity in whole paragraph, and Japanese have a tendency to pronounce vowels like schwa too loudly. From these points, it can be said that the learners are influenced by Japanese language when they pronounce schwa, but duration of schwa is the most important point (and is relatively easier to acquire).

In future, we would like to observe only subjects who had long schwa duration and see if vowel space becomes more important. There is no correlation between perceived pronunciation skill and distance of vowel to the mid-center of the vowel space in this experiment, but the long duration might change that because native listeners would notice the vowel space of schwa if it were longer. Additionally, it is better to ask native listeners to judge specifically schwa. Native listeners did not focus on schwa in this experiment. However, if they look at other sounds such as /l/ and /r/, the listeners may judge the pronunciation level not paying attention to schwa sound. Also, divide schwas into spelling and correlate with each duration in future work.
Acknowledgements

I would like to thank following people. Professor Ian Wilson was my supervisor of the research. He advised me the way to analyze, how to write the thesis, and gather the English native speakers who judged the English pronunciation skill.

And Miyuki Yoshizawa, who is another Phonetics Lab member. She supplied me with the sound files and instructed me how to make graphs in Excel.

Finally, I thank the learners and judges at the University of Aizu.

References


Appendix

The North wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak.

They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other.

Then the North wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him, and at last the North wind gave up the attempt.

Then the Sun shone out warmly, and immediately the traveller took off his cloak.

And so the North wind was obliged to confess that the Sun was the stronger of the two.