Pitch and Intensity in the speech of Japanese speakers of English: Comparison with first English speakers

Jun Okada s1170182

Abstract

The speech of L2 learners of English is often difficult to understand because of intonation problems and misplaced word stress. In this research, we investigated whether or not the intonation patterns of Japanese speakers of English show common patterns based on proficiency level. First, we recorded “The North Wind and the Sun” from 50 Japanese undergraduate students (aged 18 to 24). We recorded native English speakers and also obtained such native data online. Next, we labeled each word and analyzed the pitch and intensity using Praat. Data was separated by gender and by proficiency in English, results were plotted, and statistical analysis was undertaken. Preliminary results show that pitch (and to a lesser extent, intensity) showed a common pattern across native speakers, but that L2 speakers relied on intensity much more than pitch in the production of stress.

1 Introduction

There are two goals of this research. The first goal is to determine the relative importance of pitch and intensity to indicate stress in first (L1) and second (L2) English speech.

It has been shown that overall pitch range is greater for native speakers of English than for L2 speakers [1]. Also, as Japanese learners of English become more proficient (in listening and reading), their pitch range increases [2]. It is difficult for Japanese English learners to learn English rhythm and intonation, because of L1 and L2 (and even dialect) differences in intonation [3]. Since stress is indicated by pitch and intensity, we investigated the pitch and intensity differences between Japanese L2 English and native speakers’ L1 English.

The second goal of this research is to determine whether or not general L2 proficiency (in listening and reading) roughly correlates with pitch and intensity accuracy in L2 production. L2 general proficiency is correlated to speech sound discrimination [4], so L2 proficiency may be correlated with intonation ability of Japanese speakers’ English. We investigated whether or not the intonation patterns of Japanese speakers’ L2 English show common patterns based on general L2 proficiency level. We used TOEIC (Test of English for International Communication) to measure general language proficiency (focused on listening and reading skills), because it is taken almost students in University of Aizu. It is hoped that this research will give beneficial information to Japanese English learners to help them better their intonation and thus their communication skills.

The paragraph of “The North Wind and the Sun” was used in this research, because it was used in previous research [2] and it includes all English phonemes, and it is now used commonly in phonetics.

2 Method

2.1 Participants

In this research, all Japanese voice data was used from previous research [2]. Voice data were collected from 50 Japanese undergraduate students aged from 18 to 24 (mean = 21.0, standard deviation = 1.09). First, 43 persons’ voice data were recorded in the junior English class ”Acoustic Analysis Using Software”. Since there was not enough female data, we asked for cooperation from 7 women outside of class as well. In total, we analyzed 45 speakers’ voice data, excluding data in which audio files were not clear. Moreover, TOEIC scores were available for 37 out of those 45 people. Voice data of 3 native speakers (male 2, female 1) were downloaded from online. Voice data of 2 native speakers were recorded in the lab (both male).

2.2 Stimuli

In this research, all participants read aloud both the Japanese version and the English version of “The North Wind and the Sun” from the Aesop Fable, and we recorded each speaker to a separate sound file. The text from the Principles of the International Phonetic Association 1949, which are used broadly in research involving things such as listening tests, pronunciation practice for clinical recording, and acoustic analysis, were used in this research [5]. You can see the text in Appendix 1.
2.3 Apparatus

All recording for this research was carried out in the iLab classroom of the University of Aizu. When recording, a headset microphone (Sennheiser PC 131), an iMac (Mac OS X Version 10.6.8), and Praat (Version 5.2.38) were used. An iMac (Mac OS X Version 10.6.8) and Praat (Version 5.2.45) were used to analyze voice files, and Microsoft Excel (Microsoft for Mac 2011 version 14.1.4) was used to create tables and graphs.

2.4 Procedure

2.4.1 Stimuli Preparation

To prepare for data collection, a one-page questionnaire was constructed. Columns were made for participants to write down personal information such as student ID number, name, home prefecture, gender, age, and TOEIC score. Also, instructions for recording, such as the setup of the microphone and where to save files, were written on this paper in detail. Finally, English text of “The North Wind and the Sun” was written on the questionnaire in a large font size.

2.4.2 Data Collection

Before recording, the paper described above was distributed to all participants. They filled in all the items, and the paper was submitted after recording finished. Before recording, they read and understood the notes, and we explained again orally. Moreover, we got the participants to practice reading several times, and any words with pronunciation they did not understand were read aloud by the professor. After practicing a number of times, the participants made recordings.

2.4.3 Data Analysis

The recorded sound files were checked one by one. Because the beginning of a file and the end of a file have silent parts, they were cut and files were resaved so that it might be easy to analyze each file. All files were labeled and each syllable number was marked (First sentence 1-27, Second sentence 28-59, Third sentence 60-100). A Praat script was run in order to get pitch and intensity data (Maximum, Minimum, average pitch of the each file; standard deviation; the range for analysis in the script [male: 70-200Hz, female: 75-350 Hz]) and results were displayed on Excel files separating Japanese L2 speakers’ file and L1 speakers’ file each sentence. If the pitch of some syllables appeared much too high, we interpolated from neighboring data. All pitch data were converted to cents instead of Hertz using the formula: 1200*Log10(freq1/freq2)/Log102, where freq1 means a given speaker’s average frequency. Pitch and intensity peaks were determined by comparing a cell’s value at time = t to the values of cells at time = t ± 1; Peaks were indicated with the value “1” in the Excel table. The number of peaks was summed and indicated in the table. We determined which total pitch and intensity peaks are 1 standard deviation above overlap. And the total number were converted to percentage to know how much people stress.

3 Results

In Table 1, it can be seen that where most speakers have high pitch it the paragraph, and also in Table 2, it is indicated where most speakers have high intensity. The numbers are used Red. Thus, following table was made (J = L2 speakers, E = English L1 speakers). Low level is less than average +1 standard deviation by TOEIC, Middle level is average +/-1 stddev, and also High level is greater than average +1

---

Table 1: Percentage of number of speakers with a pitch peak on that word

|   | The | North | Wind | and | the | Sun | were | di | pu | lim | which | was | strong | ger | when | a | in | was | val | le | le | came | a | long | word | end | e | in | warm | closet |
|---|-----|-------|------|-----|-----|-----|------|----|----|----|-------|-----|--------|----|-------|---|----|----|-----|----|----|-----|----|------|-----|----|---|-----|------|
| J Low | 20 | 80 | 60 | 40 | 20 | 80 | 40 | 0 | 80 | 60 | 40 | 20 | 80 | 80 | 20 | 80 | 0 | 40 | 80 | 20 | 60 | 0 | 40 | 60 | 40 | 20 | 100 | 0 |
| J Mid | 4 | 56 | 32 | 8 | 4 | 56 | 24 | 12 | 60 | 16 | 52 | 12 | 32 | 20 | 56 | 8 | 36 | 36 | 8 | 48 | 0 | 56 | 16 | 40 | 16 | 56 | 16 |
| J High | 20 | 40 | 20 | 0 | 0 | 80 | 0 | 0 | 60 | 0 | 60 | 0 | 60 | 20 | 40 | 0 | 80 | 0 | 0 | 20 | 0 | 20 | 40 | 20 | 0 | 20 | 20 |
| E Nat | 0 | 100 | 0 | 0 | 20 | 20 | 0 | 20 | 80 | 0 | 80 | 0 | 0 | 20 | 0 | 0 | 80 | 20 | 0 | 20 | 0 | 0 | 20 | 0 | 0 | 60 |

Table 2: Percentage of number of speakers with an intensity peak on that word

|   | The | North | Wind | and | the | Sun | were | di | pu | lim | which | was | strong | ger | when | a | in | was | val | le | le | came | a | long | word | end | e | in | warm | closet |
|---|-----|-------|------|-----|-----|-----|------|----|----|----|-------|-----|--------|----|-------|---|----|----|-----|----|----|-----|----|------|-----|----|---|-----|------|
| J Low | 40 | 40 | 80 | 40 | 40 | 60 | 40 | 40 | 40 | 60 | 60 | 40 | 60 | 40 | 60 | 40 | 60 | 40 | 60 | 60 | 40 | 80 | 40 | 20 |
| J Mid | 68 | 16 | 44 | 24 | 20 | 76 | 4 | 76 | 20 | 20 | 48 | 24 | 12 | 68 | 20 | 32 | 32 | 48 | 8 | 76 | 12 | 0 | 28 | 60 | 24 | 0 |
| J High | 60 | 0 | 40 | 20 | 40 | 20 | 60 | 0 | 60 | 0 | 0 | 40 | 20 | 0 | 60 | 0 | 20 | 0 | 80 | 0 | 80 | 0 | 20 | 40 | 40 | 20 |
| E Nat | 40 | 20 | 80 | 0 | 20 | 40 | 60 | 0 | 20 | 80 | 0 | 0 | 100 | 0 | 20 | 40 | 60 | 20 | 40 | 0 | 40 | 40 | 0 | 60 | 20 | 40 | 0 |
Figure 1 and Figure 2, Figure 3, Figure 4 shows indicated that how much people were stressed in the sentence. On L2 speakers’ graphs, Led lines are high level Learners (N = 5). Green lines are middle levels (N = 25), and blue lines are low levels (N = 5).

Figure 1: Pitch of 35 Japanese L2 Speakers of English [0 cents = given speaker’s mean pitch]

Figure 2: Pitch of 5 Native Speakers of English [0 cents = given speaker’s mean pitch]

Figure 3: Intensity of 35 Japanese L2 Speakers of English [0 dB = given speaker’s mean intensity]
4 Discussion and Conclusions

For native speakers of English, pitch (rather than intensity) is a much more salient indicator of which words are stressed in a sentence. Intensity peaks are almost strange, otherwise pitch peaks is obvious correct. According to table 1 indicated pitch peaks, most native speakers have pitch peaks “North”, “pu”, “which”, “tra”, “cloak” in the sentence, but on table 2 indicated intensity peaks, they have intensity peaks “Wind”, “were”, “tra”, “in”. No speakers stressed “were”, “ting”, “in”, so pitch peaks are indicated more better about stress.

L2 English speakers tend to have many more peaks, in both intensity and pitch, than native speakers, indicating the lack of longer prosodic phrases in L2 speech. In table 1 and 2, many Japanese not depended on levels have many peaks on the sentence not only pitch but also intensity.

In the intensity results, we can see that Japanese L2 English speakers stress the first word of a phrase (“The” at the beginning of the sentence and “when” at the start of the second longer phrase).

Mid- and Low-level L2 speakers tended to stress the second-last word of the sentence though most native speakers did not.

Unlike native speakers, Japanese L2 speakers of English tended to have intensity peaks on schwa sounds, perhaps because they tend to pronounce them like Japanese [a] with the mouth wider open than a native speaker’s schwa (see “The”, “were”, “ler”, “a(long)”, and “a”).

Perhaps unsurprisingly, it appears that L2 general proficiency (as measured by reading and listening) does not roughly correlate with pitch and intensity accuracy in L2 pronunciation (production), even though it has been shown to correlate (medium strength) with L2 sound discrimination. Sound discrimination does not necessarily lead to ability to use prosody correctly in a second language.

5 Limitations and Future Research

First, automated analysis should be used to enable more data to be analyzed. It took time to check by hand and find and correct errors in pitch tracking.

Next, too few native speakers’ data were used in this research. Many more native speaker participants are needed from a variety of dialects of English.

The speaking and writing test for TOEIC and/or TOEFL may be more useful. We used the TOEIC reading and listening test to score general proficiency, but the TOEIC and TOEFL tests have writing and speaking versions too.

6 Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Professor Ian Wilson, for his guidance and support to my research work. Also, I genuinely appreciate the University of Aizu students’ participant in this study by providing valuable voice data. Thanks also to Miyuki Yoshizawa for collecting voice data.

References

Appendix 1: English stimuli used (see [5])

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.