ABSTRACT
Durational reduction of unstressed syllables is a characteristic of English rhythmic organization. This study investigated the durational patterns of unstressed vowels in the production of Japanese speakers of English, as a part of our ongoing research on prosodic transfer from L1 to L2 phonetics. The results of our experiment showed a general tendency for all four Japanese participants. On average, the duration of an unstressed vowel in a monosyllabic function word was longer than in the unstressed syllable of a polysyllabic content word (i.e. the reduction of unstressed function words is harder to learn than the reduction of unstressed syllables within content words). This pattern cannot be explained by the transfer of Japanese prosody, in which duration and word accent are independent properties. Possibly, the pattern is due to a general constraint on L2 speech development: the metrical grouping of syllables across words is harder than within a single word.

1. INTRODUCTION

1.1. Durational Reduction in English
Durational reduction of lexically unstressed syllables is one of the major characteristics of English rhythmic organization. In English, stress (word accent) influences F0, intensity, vowel quality and duration. It has been reported in earlier studies that unstressed syllables are lower in F0, weaker in intensity, more central in vowel quality and shorter in duration than stressed syllables [1, 3, 4, 6, 7, 8].

In English, the position of stress in a polysyllabic word is a lexical property. The occurrence of stress in monosyllabic words can be mostly predicted by whether that word is a content or function word. Monosyllabic content words (e.g. nouns, adjectives or verbs) are likely to be stressed. For example, the verb go is typically stressed. On the other hand, monosyllabic function words (e.g. articles, prepositions or auxiliaries) are typically unstressed except under contrastive focus. For example, in a pragmatically neutral context, the preposition to is typically pronounced as [tə], not [tu].

Thus, both unstressed syllables in polysyllabic content words and monosyllabic function words are durationally reduced. This is an important property for the English stress-timing system.

1.2. Relevant Features of Japanese
Unlike in English, duration and word accent are independent properties in Japanese. This language has a phonemic length contrast which characterizes all the vowels and most consonants (e.g. *taru* ‘catch’ vs. *taru* ‘pass’; *ka* ‘shoulder’ vs. *ka* ‘won’). The distribution of word accents hardly affects the duration of Japanese syllables, probably because the phonemic length contrast must be preserved.

Previous studies showed that the only reliable cue of the contrast between accented and unaccented syllables is F0 [4, 5, 9, 11]. Thus, Japanese does not show systematic durational reduction of unaccented syllables.

1.3. Problem in L2 English Produced by Japanese Speakers
Given the aforementioned differences in duration and word accent between English and Japanese, it is not difficult to expect that the prosodic features of L1 Japanese negatively transfer to L2 English. Indeed, difficulties in acquiring English rhythm have been observed among Japanese speakers of English.

Among other things, it is difficult for Japanese learners of English to acquire the durational contrast between stressed and unstressed syllables. Ueyama’s study [10] showed that the mastery of the stress-controlled durational contrast correlates with the oral proficiency level of Japanese learners of English.

Insufficient durational contrast between stressed and unstressed syllables in the production of Japanese speakers of L2 English may be caused by insufficient durational reduction of unstressed syllables. It has been reported that a group of non-native speakers including a Japanese speaker did not make reduced syllables significantly shorter than full syllables, while native speakers of English durationally distinguished the two syllable types [3].

1.4. Purpose of the Study
The present study is a part of our ongoing research on the prosodic transfer of L1 Japanese to L2 English phonetics. In this study, we focus on the duration of unstressed vowels in the production of L2 English by Japanese speakers. The effects of three factors on duration will be investigated. The first factor is the context of unstressed syllables: unstressed monosyllabic function words vs. unstressed syllables polysyllabic content words. The second factor is the size of Inter-Stress Interval (ISI), defined as the number of unstressed syllables between two stressed syllables: ISI=1 vs. ISI=2 vs. ISI=3. The third factor is the proficiency level of Japanese learners of L2 English.

2. EXPERIMENT

2.1. Subjects
The control group consisted of three native speakers of American English: A1, A2 and A3. A1 was a male speaker, while A2 and A3 were female speakers. The experimental group consisted of four Japanese learners of English (with various proficiency levels) who were students at UCLA at the time of the recording: J1, J2, J3 and J4. All Japanese participants spoke the standard variety of Japanese, *hyojyungo*, as their native tongue.

The English proficiency level of the Japanese participants was determined on the basis of 1) the years of residence in the United States and 2) the average percentage of time spent speaking English in a typical week. This information was collected at the time of data collection.
The proficiency ranking of the Japanese participants was determined in the following way. We first ranked the five participants on the basis of the number of years of residence in the United States. If two speakers have been in the United States for the same amount of time, we ranked them on the basis of the average percentage of time spent speaking English. The results of this rating method are presented in Table 1:

<table>
<thead>
<tr>
<th>years of residence in the US</th>
<th>J1</th>
<th>J2</th>
<th>J3</th>
<th>J4</th>
</tr>
</thead>
<tbody>
<tr>
<td>percentage of time spent speaking English per week</td>
<td>9 yrs</td>
<td>6 yrs</td>
<td>5 yrs</td>
<td>5 yrs</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>50%</td>
<td>35%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table 1. Proficiency ranking of the four Japanese participants

We are aware that other factors affect L2 proficiency development (see the references of Flege’s work). However, Table 1 provides a broad characterization of the proficiency levels of the four Japanese participants.

2.2. Speech Materials

Our corpus contained three pairs of test sentences. Sentences in each pair were identical in terms of Inter-Stress Interval (ISI), and they were different in terms of the context of the tested unstressed syllables (within a polysyllabic content word vs. in a monosyllabic function word). In order to see how the number of interstress syllables affects duration, the ISI size varied from 1 to 3. The three pairs of test sentences are listed in Table 2 (the acute mark indicates stress). We controlled the position of the tested syllables from the beginning of the sentence and syllable weight.

<table>
<thead>
<tr>
<th>ISI</th>
<th>unstressed syllable within a content word</th>
<th>monosyllabic function word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gō together</td>
<td>Sée the governor</td>
</tr>
<tr>
<td>2</td>
<td>Sée the philosophiser</td>
<td>Borrow a penny</td>
</tr>
<tr>
<td>3</td>
<td>Gō to the philosophiser</td>
<td>Hurry to the ladder</td>
</tr>
</tbody>
</table>

Table 2. Test Sentences

2.3. Procedure

The six target sentences were mixed with foil sentences and pseudo-randomized. To avoid segmental effects on duration while preserving prosodic patterns, the participants were asked to replace each syllable of the text with the syllable /no/ (reiterant speech). For example, “Go together” will be read as /NO noNO no/ (“NO” and “no” indicate stressed and unstressed syllables, respectively).

The participants were given sufficient time to practice reiterant speech ahead of recording. We made sure that all our subjects were able to perform the task. The Japanese participants were also asked to produce a set of Japanese sentences using reiterant speech. None of them showed difficulties. The speakers had to read each sentence twelve times. The first and last repetitions were not used for analysis. The data were collected in the recording booth of the UCLA Phonetics Laboratory.

2.4. Measurements

The recorded data were converted from analog to digital at a 10 kHz sampling rate. We measured the duration of the unstressed vowel /o/ for each of the six tested conditions, using Kay Elemetrics’ Computerized Speech Laboratory (CSL) hardware and software. All the measurements were based on waveform analysis and wide-band spectrograms. The energy distribution was additionally inspected in the cases of difficult segmentation. Since segmental duration is affected by the prosodic structure of utterances, we checked that each speaker produced similar intonation patterns across tokens by inspecting the sequence of pitch accents, phrase tones and boundary tones.

3. RESULTS

Vowel durations in each condition were statistically analyzed by running a series of 2-factor ANOVAs in order to examine the effects of the context of the unstressed syllable and the ISI size. Scheffe post-hoc tests were performed when necessary.

3.1. Native Speakers of English

The mean duration of the unstressed vowels! produced by the native English speakers in each context are reported in Figure 1. ANOVA results are presented in Table 3 (shaded cells represent significant effects).

<table>
<thead>
<tr>
<th>context</th>
<th>ISI</th>
<th>context*ISI</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>F(1) = .546</td>
<td>F(2) = 7.490</td>
</tr>
<tr>
<td></td>
<td>p = .4631</td>
<td>p = .013</td>
</tr>
<tr>
<td>A2</td>
<td>F(1) = 0.003</td>
<td>F(2) = 9.537</td>
</tr>
<tr>
<td></td>
<td>p = .9559</td>
<td>p = .0003</td>
</tr>
<tr>
<td>A3</td>
<td>F(1) = 3.799</td>
<td>F(2) = 8.917</td>
</tr>
<tr>
<td></td>
<td>p = .565</td>
<td>p = .0005</td>
</tr>
</tbody>
</table>

Table 3. ANOVA results for native English speakers (α = .05)

None of the three native English speakers distinguished the duration of the unstressed vowel of a monosyllabic function word from that of an unstressed syllable within a content word. This pattern was consistently observed for all ISIs.

On the other hand, the ISI size showed significant effects for all native speakers. In the speech of A1 and A2, on average, the duration of the unstressed vowel /o/ was longest for ISI=2 and shortest for ISI=3: thus, ISI=2 > ISI=1 > ISI=3. In A3’s speech, the distribution of duration on average was ISI=1 > ISI=2 > ISI=3. These patterns do not suggest any coherent tendency for the native group.

3.2. Japanese Speakers of English

Mean durations of unstressed /o/ as produced by non-native speakers are presented in Figure 2. ANOVA results are reported in Table 4 (shaded cells represent significant effects).
Figure 1a-c. Mean durations of the unstressed vowel /o/ for native English speakers

Table 3. ANOVA results for Japanese speakers of English

<table>
<thead>
<tr>
<th>Subject</th>
<th>context</th>
<th>ISI</th>
<th>context*ISI</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1</td>
<td>F(1) = 24.699</td>
<td>p = &lt; .0001</td>
<td>F(2) = 1.749</td>
</tr>
<tr>
<td>J2</td>
<td>F(1) = 18.403</td>
<td>p = &lt; .0001</td>
<td>F(2) = 6.524</td>
</tr>
<tr>
<td>J3</td>
<td>F(1) = 50.341</td>
<td>p = &lt; .0001</td>
<td>F(2) = 3.859</td>
</tr>
<tr>
<td>J4</td>
<td>F(1) = 82.707</td>
<td>p = &lt; .0001</td>
<td>F(2) = 4.983</td>
</tr>
</tbody>
</table>

Figure 2a-c. Mean durations of the unstressed vowel /o/ for Japanese speakers of English
There is a strong effect of the context of the unstressed vowel on duration for all four Japanese speakers. We additionally performed Scheffe posthoc test for each speaker to find for which ISI size there was a significant context effect (i.e. function word vs. unstressed syllable of a content word). Results are presented in Table 5 (shaded cells indicate significant differences in mean durations between the two contexts). In the production of J3 and J4, on average, vowel duration was longer in a monosyllabic function word than in an unstressed syllable of a polysyllabic content word for all ISIs. J1 showed the same pattern for ISI=2 and ISI=3, and so did J2 for ISI=1 and ISI=2.

We found significant effects of the ISI size for J2, J3 and J4. Among these three speakers, only J3 showed the same pattern within both contexts: on average, duration is longest for ISI=1 and shortest for ISI=2. The other two speakers, J2 and J4, showed significant interaction effects between the two variables. Like for the case of the native group, it seems difficult to interpret the effects of the ISI size in the non-native data.

3.3. Proficiency Effects

None of the three English speakers distinguished the mean duration of the unstressed vowel in a monosyllabic function word from the duration of the unstressed vowel in a polysyllabic content word.

The most proficient Japanese learners, J1 and J2, showed native-like patterns, but not for all ISIs. J1 showed a native-like pattern only for ISI=1, while J2 showed a native-like pattern only for ISI=3 (see Table 5). The two less proficient speakers, J3 and J4, showed non-native patterns for all ISIs: i.e. the duration of unstressed /o/ was consistently longer in a monosyllabic function word than in a polysyllabic content word. Thus, it seems that chances to neutralize the duration of unstressed vowels within both contexts correlate with proficiency levels at least to some extent.

4. DISCUSSION

The data show a general pattern across the Japanese learners: on average, the duration of unstressed /o/ in a monosyllabic function word is longer than in a polysyllabic content word. Thus, the reduction of unstressed function words is harder to learn than the reduction of unstressed syllables within a content word.

This tendency cannot be directly predicted by the prosodic transfer of L1 Japanese. Give that duration and word accent are independent properties in Japanese, we may not expect durational difference between the two unstressed syllable contexts in L2 English produced by Japanese speakers. Then, where does the systematic non-native pattern come from? Possibly, the reduction of function words is harder because the metrical grouping of syllables across words is more difficult than within a single word. Thus, non-native speakers tend not to treat function words as phonological clitics, but as independent prosodic units, which cannot be entirely stressless (and hence reduced). Further investigation is needed to assess the validity of this hypothesis.

5. Conclusion

Durational reduction of unstressed syllables is a characteristic of English stress-timing. This study investigated how Japanese speakers of English produced English unstressed syllables. The results of our experiment show that unstressed vowel reduction is harder in monosyllabic function words than in content words. This pattern cannot be predicted by the direct transfer of L1 Japanese characteristics. A possible explanation relates this pattern to a general constraint on L2 speech development: the metrical grouping of syllables across words is harder than within a single word.

Notes

1. On-going research shows that vowels labeled here as “unstressed” are consistently shorter than the corresponding stressed vowels.

REFERENCES