Dichotomy or Trichotomy?
Testing the Perceptual Boundary of Quantity
between Japanese and Finnish

Toshiko Isei-Jaakkola

Department of Phonetics
University of Helsinki, Finland
Toshiko.Jaakkola@helsinki.fi

Abstract
The results of two kinds of perception test on the perceptual boundary ranges of quantity distinctions between Japanese and Finnish were compared. Test A was conducted based on the trichotomy concept, choosing one from short, long and uncertain. Test B, based on the dichotomy concept, was conducted by choosing either short or long, utilising the same test words. The materials were 120 discrimination test sets at the word level designed to study the perceptual boundaries of short/long quantitative distinctions of vowels/consonants. The bisyllabic synthetic nonsense words used had eight kinds of syllable structure and five different prosodic conditions with three kinds of pitch and intensity variance pattern only for the vowels. The phonemes were /a, m, p, s/. Both tests, done under the same conditions, utilised the same number of Japanese and Finnish subjects (n = 7 x 2). The response number was 47,040, and the number of words to be analysed 3,360. The results showed significant differences within the segment and word between the two tests, the perceptual boundaries being shorter in Test B than in Test A for vowels and consonants. The standard deviations and variation ranges within the segment and word were also smaller in Test B than in Test A. There were no significant differences between the two languages in Test B. These findings may shed some light on the testing methodology in the perception tests.

1. Introduction
Japanese and Finnish can be compared in terms of the durational differences between linguistically distinctive short/long vowels and short/long consonants.

Perceptual identification tests do not often consider syllable structures and varying F0 and intensity level, and the data has been analysed by the binary concept of quantity, short or long for both languages. For example, the boundary range problem has been discussed by Lehtonen [1] for Finnish and Fujisaki and Sugitou [2] for Japanese, etc., but such studies have been much fewer, compared with the study of quantity in production.

The main point of this study was to compare the perceptual boundary ranges between a trichotomy, i.e., three response choices (short, long or uncertain) [3], and a dichotomy (short or long), i.e., the binary concept.

The perceptual boundary range means the durational border between the two quantity categories (V/VV, C/CC) and indicates the auditory uncertainty area (U) between short (S) and long (L) segments, as Figure 1 illustrates.

![Figure 1: Illustration of the perceptual boundary area lying between short and long segments. S = short, L = long, and U = uncertain.](image)

In this study (Test B), only the testing methodology has been changed, utilising exactly the same materials as in Test A in order to compare these two test results.

The following questions arose:

(1) Is there a difference between the test results based on the trichotomy and dichotomy concepts?
(2) Is there a difference between Japanese and Finnish in terms of perceptual boundary ranges, comparing these test results?
(3) To what degree do such prosodic variants as F0 and dB, different syllable structures, and variable word durations affect perceptual boundaries in Finnish and Japanese within the segment and word, comparing these test results?

2. Experimental procedure
For the perception tests, I used synthetic bisyllabic nonsense words in isolation. Stimuli were produced using an Infovox speech synthesizer simulating a male voice.

2.1. Experimental designs
There were three experimental designs. The first was to create four different kinds of bisyllabic word structures for vowels – CVCV, CVCCV, CVVCV, CVVCVV and
four for consonants – CVCCV, CVCCVV, CVVCCV, CVVCCVV, all of which exist in Finnish and Japanese (except for CVVCCVV). These structures include from two to five morae according to Japanese mora-counting.

The second design was to create stimulus words of different word lengths. The phonemes used were /a, m, p, s/. The number of word-medial and word-final vowel stimuli continua (50-200 ms) was 16 and the number of corresponding consonants for word-medial /p – pp/ (90-200 ms) was 12, 11 for /m – mm/ (60-160 ms) and 13 for /s – ss/ (80-200 ms) with a 10 ms incremental increase for all four phonemes. The word duration thus varied from 220 ms – 690 ms.

Table 1: Five prosodic conditions and acoustic parameters (pitch and amplitude) in three patterns.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Acoustic Parameter</th>
<th>1st syllable</th>
<th>2nd syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level F0</td>
<td>Pitch</td>
<td>100Hz</td>
<td>100Hz</td>
</tr>
<tr>
<td>Level dB</td>
<td>Amplitude</td>
<td>26dB</td>
<td>26dB</td>
</tr>
<tr>
<td>HL</td>
<td>Pitch</td>
<td>120Hz</td>
<td>95Hz</td>
</tr>
<tr>
<td>Level dB</td>
<td>Amplitude</td>
<td>26dB</td>
<td>26dB</td>
</tr>
<tr>
<td>LH</td>
<td>Pitch</td>
<td>95Hz</td>
<td>120Hz</td>
</tr>
<tr>
<td>Level dB</td>
<td>Amplitude</td>
<td>26dB</td>
<td>26dB</td>
</tr>
<tr>
<td>SW</td>
<td>Pitch</td>
<td>100Hz</td>
<td>100Hz</td>
</tr>
<tr>
<td>Level F0</td>
<td>Amplitude</td>
<td>29dB</td>
<td>26dB</td>
</tr>
<tr>
<td>WS</td>
<td>Pitch</td>
<td>100Hz</td>
<td>100Hz</td>
</tr>
<tr>
<td>Level F0</td>
<td>Amplitude</td>
<td>26dB</td>
<td>29dB</td>
</tr>
</tbody>
</table>

Furthermore, prosodic conditions were added to those above. There were five patterns in the prosodic conditions as in Table 1. The label ‘Level’ for prosodic conditions represents unchanged parameters in the stimulus word. ‘HL’ represents the pitch in the first syllable being higher than that in the second syllable but intensity in the word remaining unchanged. ‘LH’ represents lower pitch in the first syllable than in the second but unchanged intensity. ‘SW’ represents greater intensity in the first syllable than in the second but unchanged pitch. ‘WS’ represents lower intensity in the first syllable than in the second but unchanged pitch. Only the vowel in the first and second syllables had unchanged F0 – 100 Hz and changed F0 – 95 Hz/120 Hz, 120 Hz/95 Hz, and unchanged intensity – 26 dB and changed intensity – 26/29 dB, 29/26 dB. The following details were confirmed before the experiments: (1) 25 Hz (4.0441 semitone differences) was enough to differentiate higher or lower pitch between syllables, and (2) 3 dB was also sufficient to differentiate the syllable with higher or lower intensity. Hence, there were five prosodic conditions yielding complex stimuli in eight different syllable structures, and all variants add up to 120 test sets (8 SS (= syllable structures) x 5 PC (= prosodic conditions) x 3 Cs (= kinds of consonants)) for vowel and consonant distinctions.

2.2. Test methods and subjects

In Test B, seven Japanese subjects who are Tokyo dialect speakers and seven Finnish speakers from Helsinki and its surroundings participated in the discrimination tests and were asked to choose one out of two choices (short or long). These responses were categorised into short, long or uncertain. The boundary range values were thus acquired from the uncertainty area, which lies between 100% for short and 100% for long.

Thus, the number of word responses was 1,680 (N14 x 120 test sets) and the word responses 23,520 for Test B. Since this response number was the same in the tripartite choices (Test A), the number of word responses to analyse was 3,360 (47,040 segment responses) altogether.

3. Results and analyses

The overall mean values of the boundary ranges, variation ranges (the difference between maximum and minimum value) and standard deviations for all responses were compared between the vowels and consonants of each language in different syllable structures and variable prosodic conditions.

3.1. Distribution of response time

As Figure 2 illustrates, the response time in Test A stretched maximally 60 ms in Finnish and 50 ms in Japanese and concentrated mainly from 10 to 30 ms, whereas in Test B it stretched maximally 50 ms for Japanese and 30 ms for Finnish and approximately 99% of the response time concentrated within 10 ms in both...
languages. A significant difference between Test A and Test B in terms of boundary range emerged.

3.2. Overall mean boundary ranges

The difference in overall mean boundary range durations between Finnish and Japanese in Test A was 3.15 ms (Finish 19.09 ms > Japanese 15.94 ms). On the other hand, that in Test B was only 0.04 ms (Finnish 10.21 ms < Japanese 10.25 ms).

In three choices (Test A), the overall mean durational values of the boundary range (BR) for the vowels were longer in Finnish (19.9 ms) than in Japanese (15.8 ms). The same values for the consonants were also longer in Finnish (18 ms) than in Japanese (16.1 ms), demonstrating that the overall boundary ranges are longer in Finnish than in Japanese. In other words, the Japanese subjects made their decision in less time than their Finnish counterparts in differentiating between short and long when they were given three choices.

On the other hand, given two choices, BR values were slightly shorter in Finnish (10.1 ms) than in Japanese (10.2 ms) for vowels and slightly shorter in Japanese (10.25 ms) than in Finnish (10.31 ms) for consonants, values which indicate that the BR differences were not significant between the two languages when the speakers were given two choices. Thus the explicit difference was that of BR values between Test A (three choices) and Test B (two choices), as Figure 3 illustrates.

Table 2 shows the variation ranges (VR) and standard deviation (SD) between Finnish and Japanese for vowels and consonants. Table 2: The overall variation range values (VR, ms) and standard deviation values (SD, ms) of the boundary ranges between Finnish and Japanese for vowels and consonants.

<table>
<thead>
<tr>
<th></th>
<th>F2</th>
<th>F3</th>
<th>J2</th>
<th>J3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VR</td>
<td>0.6</td>
<td>4.6</td>
<td>1.2</td>
<td>3.4</td>
</tr>
<tr>
<td>SD</td>
<td>0.3</td>
<td>3.9</td>
<td>0.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Consonants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VR</td>
<td>1.7</td>
<td>4.2</td>
<td>0.0</td>
<td>2.9</td>
</tr>
<tr>
<td>SD</td>
<td>0.9</td>
<td>1.8</td>
<td>0.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

3.3. Boundary ranges in different syllable structures

Table 2 shows the variation ranges (VR) and standard deviation (SD) between Finnish and Japanese for vowels and consonants. In Test A (three choices) the Finnish speakers had longer BRs, greater SDs and wider VRs than their Japanese counterparts for both vowels and consonants. In Test B, this pattern was similar for the Finnish consonants but not for the Finnish vowels.
segment, W=within the word. 1=CVCV, 2=CVCVV, 3=CVVCV, 4=CVVCVV for vowels, 1=CVCCV, 2=CVCCVV, 3=CVVCCV, 4=CVVCCVV for consonants.

In terms of BR values (ms) and BR ratios (%), the BR difference between Finnish and Japanese in Test A was clear (F > J), but in Test B there was no significant difference between them within the segment and word.

3.4. Boundary ranges under different prosodic conditions

Figure 6 illustrates the BR values (ms) in five prosodic conditions for vowels and consonants respectively between Japanese and Finnish. Figure 7 illustrates the overall BR ratios between Finnish and Japanese for vowels and consonants within the segment (S) and word (W) under different prosodic conditions.

In terms of BR ratios, the difference between Finnish and Japanese in Test A was relatively clear (F > J), but in Test B there was no significant difference between the two languages within the segment and word.

Comparing the segment and word durational values and ratios between in different word structures and prosodic conditions, the BRs were shorter under prosodic conditions than in different word structures. The BRs of the vowels were longer than those of consonants in Test A, but in Test B the BR ratios of the consonants within segments were relatively high compared to vowels.

4. Conclusions

In Test A [3], the results showed that Finnish had longer boundary range durations and durational ratios than Japanese for both vowels and consonants within the segment and word, the effect of the different syllable structures and prosodic variants showed similar patterns in both Finnish and Japanese, and Finnish had wider variation ranges but smaller variants in terms of boundary range than Japanese. It was made clear that word structural difference (syllable structures) had more effect than the prosodic conditional differences in differentiating between short and long segments in both Finnish and Japanese, the vowels being more affected than consonants in both cases.

The following findings were confirmed by the above observations: (1) the BRs in two choices were much shorter in the duration and lower in the ratios in both Japanese and Finnish than in three choices, (2) the word-structural differences had more effect than the prosodic conditional differences in differentiating between short and long segments in both Finnish and Japanese in both Tests A and B, but (3) the BR consonantal ratios were relatively high compared to those of vowels in two choices. (4) BRs showed no significant difference between the languages in Test B (two choices). These findings may shed some light on the testing methodology in the perception tests.

5. References