

# Durational Variability of Vowel Quantity Boundary for Japanese, Finnish and Czech Speakers in Perception

Toshiko Isei-Jaakkola

Department of English Language and Culture, Chubu University, Japan

tiseij@isc.chubu.ac.jp

## Abstract

The discrimination tests were conducted at the word level to study the durational variability of vowel quantity boundary in perception, utilising disyllabic synthetic nonsense words. Four kinds of word structures and five kinds of pitch and intensity variance patterns were used. The number of the tests became 60. 21 Japanese, Finnish and Czech speakers participated in these tests as the subjects. The results showed that the overall durations of the perceptual boundary range was longest in Finnish; the count concentrated in a shorter time in Japanese and Czech than in Finnish; in relation to word structures and prosodic conditions, Finnish took the longest time in all four structures; in the durational ratios within a segment and word, Finnish and Czech showed the similar ratios according to the word structures within a word; the Finnish were influenced most of all the languages by all consonants; CVVCV – CVVCV affected Finnish and Czech most of all word structures and prosodic conditions.

**Index Terms:** durational variability, vowel quantity, pitch, intensity, word structures, different consonants, Japanese, Finnish and Czech speakers

## 1. Introduction

I define the perceptual boundary range as the uncertain area in differentiating a short segment from a long segment. The perceptual boundary range indicates the uncertain part lying between categorically short and long segments in perception (see Fig. 1).



S = short segment, L = long segment, and U = uncertain.

Figure 1: An illustration of perceptual boundary range (U) in perceptual categorisation for short and long segments. The dotted line shows PSE.

Perceptual identification tests do not often consider syllable structures and varying  $f_0$  and intensity level, and the data have usually been analysed by a binary concept of quantity, short or long (e.g., [1], [2], etc.).

Japanese and Finnish have quantity distinction between both short and long vowels and consonants ([3]), while Czech has these distinctions only between short and long vowels ([4]). These are categorically distinctive in phonetic duration, and phonologically and linguistically (semantically) distinctive. Czech has quantity distinction among Slavic languages as well as Slovak. Japanese and Czech have five vowels, while Finnish has eight. Japanese is a mora-counting and pitch-accented language, while Finnish has both trochaic stress-rhythm by syllable, and Czech is usually grouped among the syllable-timed language. Japanese can have the syllable

structure (C)V(C), if the syllable concept is adopted. Finnish has eight kinds of syllable structures, allowing the structure (C)V(V)(CC) to be valid. The Czech syllable structure shows greater variability, creating even a complicated structure such as (CCCC)V(V)(CCC).

The experiments will be conducted to investigate whether and how pitch and intensity variants and four different word structures affect the durational perceptual boundary ranges of Japanese, Finnish and Czech speakers at the word level. The four disyllabic word structures are: CVCV, CVCVV, CVVCV, and CVVCVV. These three languages have disyllabic real words with the same word structures. Because none of these three languages share meaningful words in these combinations with exactly the same short and long vowels and the same consonants, nonsense words with the combinations of C: /m, p, s/ and V: /a/ will be used for this study. The /a/ in Japanese, Finnish and Japanese is a back vowel (the Finnish one is slightly more backish) and so the quality is almost similar (cf. [3]). And it seems that the vowel quality /a/ does not differ between short and long ones, unlike in Hungarian.

In this study the following questions will be investigated to show whether there are differences or not between these language speakers.

- (1) How can perceptual boundary duration of vowel quantity differ when the pitch is changed?
- (2) How can perceptual boundary duration of vowel quantity differ when the intensity is changed?
- (3) How can perceptual boundary duration of vowel quantity differ when the word structure is changed?
- (4) How can perceptual boundary duration of vowel quantity differ when the consonant is changed?

## 2. Experimental procedures

Stimuli were created from disyllabic nonsense words based on the above syllabic structures. Stimuli were produced using an Infovox speech synthesizer. V was always /a/ and the alternative Cs were /p, m, s/ in synthetic words like /papa, mama, sasa/, etc. The number of vowel stimuli (50-200 ms) was 16 with /p/ (90 ms), /m/ (60 ms) and /s/ (90 ms for the word-initial position; 80 ms for the word-medial). Only vowels had a 10 ms incremental increase in every word (the underlined part in Table 1). The word durations are listed in Table 1. In addition, the first syllable and the second syllable had unchanged  $F_0$  (100 Hz→Level) and changed  $F_0$  (95 Hz/120 Hz→HL, 120 Hz/95 Hz→LH), and unchanged intensity (26 dB→Level) and changed intensity (26/29 dB→SW, 29/26 dB→WS). Hence, there were five kinds of conditions in four different word structures and all this adds up to 60 test sets for vowel distinctions.

Seven Japanese subjects who are Tokyo dialect speakers, seven Finnish speakers from Helsinki and its surroundings and seven Bohemian dialect speakers participated in the

discrimination tests and were asked to choose one out of three choices including 'uncertain'. Thus, the number of word tokens became 1,248 (19,968 word responses) after reducing the number of no responses (12).

In the calculation, the mean durational values of the perceptual boundary range were acquired from the uncertainty area, which lies between 100% for short and 100% for long.

Table 1. The durations of each stimulus word.

	Word Structure	C=m	C=p	C=s
1	CVCV-CVCVV	220-370	280-430	270-420
2	CVCV-CVVCV	220-370	280-430	270-420
3	CVCVV-CVVCVV	370-520	430-580	420-570
4	CVVCV-CVVCVV	370-520	430-580	420-570

### 3. Results and analysis

The overall mean values of the perceptual boundary ranges for all responses were compared between the three languages, based on different word structures and various prosodic conditions. Conclusions.

#### 3.1. Distribution of perceptual boundary ranges

##### 3.1.1. Overall results

The overall mean values of the perceptual boundary range were 15.8 ms for Japanese, 19.9 ms for Finnish and 16.3 ms for Czech (see Fig. 2). These values show that the Japanese subjects took the shortest time of all in differentiation between short and long vowels and the Finnish the longest time. The overall mean SD for Finnish was 3.8 ms, 0.8 ms for Japanese and 2.5 ms for Czech. The overall mean range (*R*) for Finnish was 3.8 ms, 0.8 ms for Japanese and 2.5 ms for Czech.

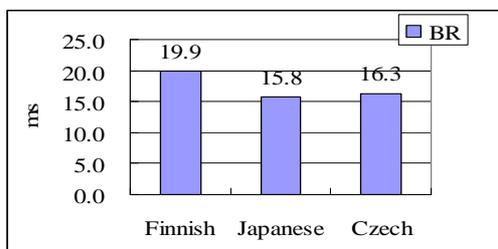


Figure 2: The overall mean values of the perceptual boundary range.

##### 3.1.2. Count

Figure 3 illustrates the number of response times. In all three languages, the highest count was 10 ms. The longer the perceptual boundary range became the count decreased in three languages. The Czech and Japanese count showed the similar decreasing patterns, being considerably dense with 10 ms and 20 ms whereas the Finnish count was gradually decreasing from 10 ms to 40 ms.

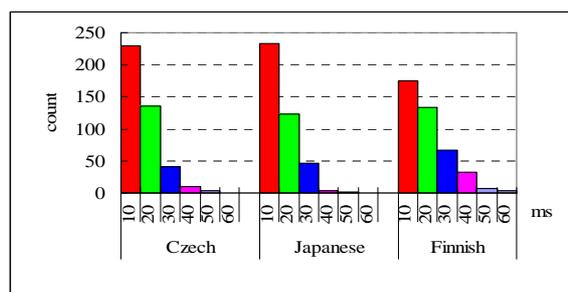


Figure 3: Count for the frequencies of perceptual boundary ranges for vowels. The number on the horizontal scale at the bottom shows each response time (ms). Count totaled 1,248: Czech= 420, Japanese=409, Finnish =419.

#### 3.2. Conditional differences: word structures vs. prosodic conditions

##### 3.2.1. Mean

The overall mean durations of the perceptual boundary range according to the word structures and prosodic conditions were calculated and translated into Figure 4. The Finnish subjects perceived vowel differentiations within the longest perceptual boundary range of all in all four structures and five prosodic conditions. In Finnish, also, the perceptual boundary range varied much more depending on word structures and prosodic conditions than the other two languages. The distributions of the perceptual boundary range between Japanese and Czech look similar, except for Czech in CVVCV-CVVCVV (19.9 ms). The longest perceptual boundary range was Finnish CVVCV-CVVCVV (25.3 ms) structure.

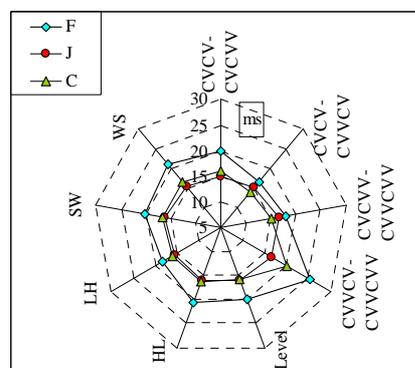


Figure 4: The distribution of perceptual boundary ranges of vowels for Japanese, Finnish and Czech according to word structures and prosodic conditions. J=Japanese, F=Finnish, C=Czech.

Comparing the durational differences between the perceptual boundary range either by word structures or by prosodic conditions, thus, it is clear that the word structures affected the perceptual boundary range more than the prosodic conditions did, particularly in Finnish and Czech.

### 3.2.2. SD

The overall mean durations of the SD according to the word structure and prosodic conditions were calculated and translated into Figure 5. Czech had the highest SD of all in all four structures. Japanese had the lowest in all four structures and all five prosodic conditions. The longest SD was in Czech CVVCY – CVVCVY structure word (10.3 ms). The shortest was in Japanese HL (0.5 ms). The order of the distributions of the SD according to the word structures was from the lowest to highest: Japanese (1.2 ms) < Finnish (1.8ms) < Czech (7.3 ms). The same according to the prosodic conditions was: Japanese (1.4 ms) < Finnish (4.0 ms) < Czech (7.3 ms). It implies that in terms of the variance the Finnish was more affected by prosodic conditions than word structures. It was not the case for Czech SD. Japanese SD shows that the Japanese subjects were hardly affected by either word structures or prosodic conditions.

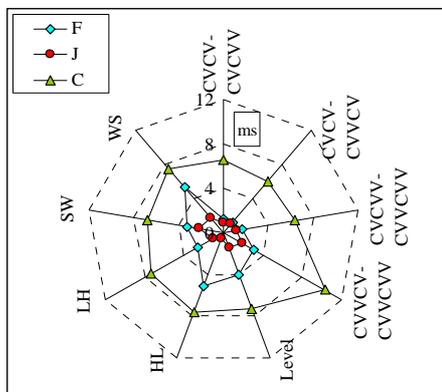


Figure 5: Distributions of the SD of the perceptual boundary ranges of vowels for Japanese, Finnish and Czech according to word structures and prosodic conditions.

### 3.2.3. Comparison between word structures and prosodic conditions

Comparing the durational differences between the perceptual boundary range either by word structures or by prosodic conditions, thus, the word structures affected the perceptual boundary range more than the prosodic conditions did. And the word structure CVVCY – CVVCVY affected the subjects' judgement in all three language most of all. Finnish was affected by word structures most of all the languages, but Czech had the highest SD, i.e., variance, of all.

### 3.2.4. Within a segment and word

These overall mean durations of the perceptual boundary range according to the word structures were calculated in percentage within a segment and word according to the word structures. The ratios of the perceptual boundary ranges in Figure 4 were calculated on the basis of 100 % for segments and 100% for words respectively. The results are listed in Table 2.

Table 2. Perceptual boundary range ratios for Japanese, Finnish and Czech according to word structures within a segment and word (in percentage).

	segment			word		
	F	J	C	F	J	C
CVCV-CVCVV	10	8	11	5	4	4
CVCV-CVVCV	8	8	9	4	4	4
CVCV-CVVCV	9	8	10	3	3	3
CVVCY-CVVCVY	13	8	13	5	3	4

Finnish and Czech showed the similar ratios according to the word structures within a word, but it was not the case for the Japanese that showed very stable SD regardless the word structures. However, these patterns became similar within the word in all these three languages.

### 3.3. Influence by surrounding consonants

#### 3.3.1. Influence by word structures

Figure 6 illustrate the mean durations of the perceptual boundary ranges for vowels according to the surrounding consonants for Japanese, Finnish and Czech in four word structures. In any consonantal environments (/m, p, s/), the Finnish were influenced most of all the languages in CVVCY – CVVCVY structure word (< 25 ms). For the Czech, only /s/ was the case (< 23 ms) in the same (CVVCY – CVVCVY) structure word. The Japanese had stable perceptual boundary range durations under these environments. The order from the longest to the shortest perceptual boundary range duration according to the consonants is as follows:

/m/ F 19.6 ms > Japanese 16.6 ms > Czech 15.1 ms,  
 /p/ F 19.3 ms > Czech 15.3 ms > Japanese 15.1 ms,  
 /s/ F 20.9 ms > Czech 18.4 ms > Japanese 15.7 ms.

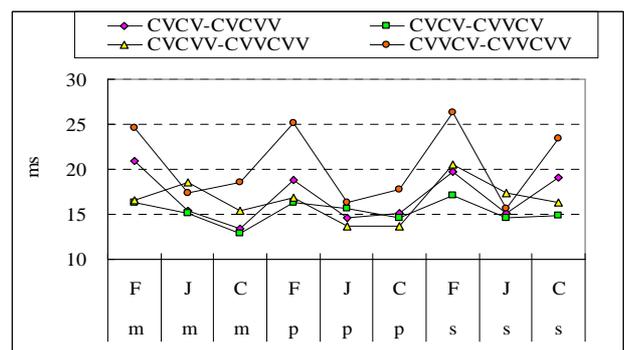


Figure 6: The mean durations of the perceptual boundary range according to the surrounding consonants for Japanese, Finnish and Czech in four word structures. /m, p, s/ are surrounding consonants.

The overall SD was highest in Finnish (3.84 ms) and smallest in Japanese (0.82 ms). (2.54 ms for Czech) The order

of the SD from the lowest to the highest according to the consonants was:

/m/ Finnish (3.9 ms) > Czech (2.6 ms) > Japanese (1.6 ms),  
 /p/ Finnish (4.6 ms) > Czech (1.7 ms) > Japanese (1.2 ms),  
 /s/ Finnish (3.9 ms) > Czech (3.8 ms) > Japanese (1.2 ms).

Therefore, Finnish vowels were affected by all three kinds of surrounding consonants used for the tests most of all.

### 3.3.2. Influence by prosodic conditions

Figure 7 illustrate the mean durations of the perceptual boundary range according to the surrounding consonants for Japanese, Finnish and Czech in the five prosodic conditions. The Finnish had the longest overall mean durations of the perceptual boundary range in all surrounding consonants of all three languages, particularly in HL (24.3 ms) surrounded by /s/. On the other hand, Finnish had the shortest duration (12.9 ms) when surrounded by /s/. The order of the mean perceptual boundary range from the longest to the shortest is as follows:

/m/ F 20.9 ms > Japanese 15.4 ms > Czech 15.1 ms,  
 /p/ F 18.9 ms > Czech 15.3 ms > Japanese 14.6 ms,  
 /s/ F 19.7 ms > Czech 18.4 ms > Japanese 15.1 ms.

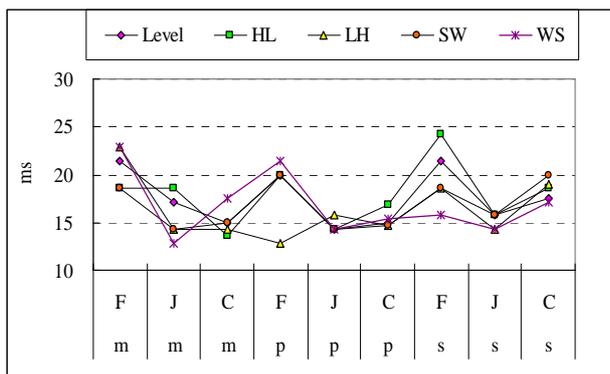


Figure 7: The mean durations of the perceptual boundary range according to the surrounding consonants for Japanese, Finnish and Czech in five prosodic conditions.

The overall SD was highest in Japanese (1.0 ms) and smallest in Czech 0.4 ms. (0.8 ms for Finnish) The order of the SD from the lowest to the highest was: Japanese (2.3 ms) < Finnish (2.2 ms) < Czech (1.5 ms) for /m/, Finnish (3.4 ms) < Czech (0.9 ms) < Japanese (0.6 ms) for /p/, and Finnish (3.3 ms) < Czech (1.1 ms) < Japanese (0.8 ms) for /s/. Japanese was affected by /m/ much more than /p/ and /s/.

## 4. Conclusions

From the above observations, I can draw the following conclusions: (1) in terms of the duration of the perceptual boundary range, the overall durations of the perceptual boundary range was longest in Finnish; the count concentrated in a shorter time in Japanese and Czech than in Finnish; (2) in relation to word structures and prosodic conditions, Finnish took the longest time in all four structures; CVVCV – CVVCVV structure affected the perceptual boundary range most of all word structures and prosodic conditions more in

Finnish and Czech than Japanese; however, Czech had the highest SD; Finnish had higher SD in the prosodic conditional differences than in word structure differences, (3) in the durational ratios within a segment and word, Finnish and Czech showed the similar ratios according to the word structures within a word, but it was not the case for the Japanese showing very stable perceptual boundary range; however, these patterns became similar within the word in all these three languages, (4) in relation to the consonantal environments, the Finnish were influenced most of all the languages by all consonants; of the word structural differences, CVVCV – CVVCVV affected Finnish and Czech most of all; the influence by different consonants according to the word structural differences was longest in Finnish as well as its SD, and so as according to prosodic conditional differences. But Japanese SD was the highest of all prosodic conditions.

These perception tests clarified that quantity between short and long vowels are categorically distinctive in phonetic duration on perception under the conditions of different word structures and prosodic variants as well as in production (see [4]) for these language speakers. Also, the perceptual durational boundaries in the three languages were not contradictory to JND (10 and 40 ms in [5]), although it denotes in production.

These results might be able to be considered in the phonological theory, particularly in Czech, different from its tradition. Also, the results from perception and production tests that I so far have shown on quantity may shed light on not only the linguistic timing issues but also on the relationship between quantity and phonotactics in respective language.

For future studies, I would like to compare the perceptual boundary range in these tripartite choices with in binary choices utilising the three languages I used for this study. Also, the relationship between the results from this study in perception and the results in production ([4]) should be compared in all these three languages. For example, it must be investigated why CVVCV – CVVCVV structure is more related in Finnish and Czech but not in Japanese.

## 5. References

- Fujisaki, H. & M. Sugitou. (1977). *Onsei no Butsuriteki Seishitsu* [The Physical Nature of Sound], Iwanami kouza *Nihongo 5 On'in* [Japanese vol. 5 Phonology]. Tokyo: Iwanami Shoten Pub. Co.
- Aoyama, K. (2001). *A Psycholinguistic Perspective on Finnish and Japanese Prosody: Perception, Production and Child Acquisition of Consonantal Quantity Distinctions*. Boston, MA: Kluwer Academic Publishers.
- Isei-Jaakkola, T., "Quantity Comparison of Japanese and Finnish in Various Word Structures", *Proceedings of International Congress EUROSPEECH Geneva*, 85-88, 2003.
- Isei-Jaakkola, T., and S. Asakawa. "A Cross-linguistic Study of Vowel Quantity in Different Word Structures: Japanese, Finnish and Czech", *Proceedings of International Conference of EUROSPEECH 2005 Lisbon*. 2905-2908, 2005.
- Lehiste, L., *Suprasegmentals*. Cambridge, MA: MIT Press, 1970.