PERCEPTUAL DEVELOPMENT ON THE IDENTIFICATION OF LENGTH IN L2 JAPANESE

Izumi Takiguchi

Graduate School of Foreign Studies, Sophia University, Japan / JSPS Research Fellow
i_takiguchi@sophia.ac.jp

ABSTRACT

This study explored perceptual development on the identification of word-final vowel length in L2 Japanese in terms of boundary location and width. Native speakers of Chinese (NC) and Japanese (NJ) participated in the study. The NC were further classified into two, NC-higher or NC-lower.

NC-lower demonstrated wider boundary width and greater boundary location than the other two did, but NJ and NC-higher were not significantly different. These results suggest that L2 learners' perception function in the L2 can approximate to native speakers' as learning proceeds. In addition, pitch affected both NC-higher's and NC-lower's boundary location in the same manner as it did to NJ's. NC may have learned such perception; moreover they may have modified their perception in accordance with NJ's production.

Keywords: L2 speech perception, Japanese vowel length, L1 Chinese, boundary location and width

1. INTRODUCTION

Japanese is one of the languages exhibiting length contrast, which is primarily cued by duration [2]. Many studies (e.g. [10]) have pointed out that acquiring such contrasts is one of the most challenging areas for second language (L2) learners, especially in the word-final position (e.g. [6]). L2 learners, whose first language (L1) does not exploit length distinction, own one category on a durational continuum for their L1 but they need to have two categories by setting a boundary on the continuum when trying to learn Japanese. Moreover, they are expected to have similar boundary location and boundary width as NJ do for the successful acquisition.

In addition to length contrasts, Japanese has a lexical pitch accent. Recent studies ([3], 5, [9]) have indicated that pitch functions as the secondary cue for NJ's length identification and a falling tone within a vowel increases their perceived vowel duration. [5] states that Japanese phonology, in which a falling tone can occur only on a long vowel, triggers this increase.

[8] explored the effects of pitch cues on the boundary of vowel length in L2 Japanese. It was shown that responses as a long vowel exhibited by NJ and NC increased as function of vowel duration, that NC's boundary values were a little greater but not significantly different from NJ's, that a falling tone affected NC's perception in the same manner as it does to NJ's. She thinks that NC's perception function can approximate to NJ's but her study did not analyze the development.

In [11] stimuli were presented to NJ and NC in ascending or descending order. Based on the results that beginners had greater values than NJ when stimuli were presented in ascending order but smaller values when presented in descending order, the author says that beginners can identify length correctly only when duration is very short or long. Advanced learners demonstrated greater values in descending than in ascending order, which was reverse shown by NJ and beginners; therefore, he proposes that they use their internal standards for identification and do not have a boundary. He concludes that NC's problem is not to set a boundary but categorical perception itself.

[4] reported that NC's boundary values were smaller than NJ's, which were similar among the learners at three levels. Besides, she showed that NC's judgment was less categorical and stable than NJ's, but advanced learners were better in the two aspects than beginners were. Consequently, she claims that L2 learners are able to own a boundary and able to come to perceive length categorically.

While [8] states that NC's boundary value can be similar and their perception function approximates to NJ's, it is said that NC may not have a boundary [11] or their boundary location does not change over time [4]. Therefore, it remains unclear how far NC can approximate their perception function to NJ's, whether they own a boundary and how their perceptual ability develops. The primary purpose of this study is to explore NC's perceptual development on the identification of word-final vowel length in Japanese in terms of their boundary location and boundary width with the consideration of pitch effects.

2. METHODS

2.1. Participants

NJ and NC participated in this study. None reported any hearing problems. There were nine participants in the NJ group, 1 male and 8 females, who speak Tokyo Japanese. The mean age of the NJ group was 37.0 (19-61). The NC group consisted of 12 participants, 1 male and 11 females. The mean age of the NC group was 27.9 (22-41).

Prior to the current study all the NC had participated in the two experiments, in which they identified final vowel length of real and nonsense words (natural stimuli). Based on an average percentage of correct answers (% correct) of the two experiments, each participant was classified as a member of NC-higher or NC-lower group as in Table 1. Mean % correct of the NC-higher (91.9%) was significantly higher than that of the NC-lower group (77.0%) [t =-6.1, df =10, p< .001].

Table 1: Individual % correct and classification of NC.

	% cor	group			% cor	group
NC01	97.3%	H		NC07	80.4%	L
NC02	99.3%	H		NC08	76.6%	L
NC03	92.7%	H		NC09	76.9%	L
NC04	88.3%	H		NC10	75.1%	L
NC05	88.9%	H		NC11	75.0%	L
NC06	84.7%	H		NC12	78.2%	L
Mean	91.9%		>***	Mean	77.0%	

2.2. Sound stimuli

All stimuli were created from a token of a nonsense word /mamama/ with HHH (H=high pitch) accent pattern produced by a male Japanese native speaker in his twenties, who is a researcher in the field of phonetics and phonology. He uttered it three times at his normal speaking rate in isolation. They were recorded using a linear PCM recorder (SONY PCM-D1) at a 44.1 kHz sampling rate and 16-bit quantization. A token with the most similar F0 values on the three syllables was selected. Table 2 shows its segment durations and F0 values.

Table 2: Duration and F0 of the original token.

segments	m	a	m	a	m	a
duration (ms)	35.1	100.4	63.6	124.8	76.5	153.1
F0 (Hz)	156.1		156.5		155.2	

A total of 44 stimuli (11 durations × 4 pitch patterns) were created by manipulating the final vowel duration and the pitch contour of the original token with Praat [1]. First, duration of the final vowel was edited by deleting or copying a pitch period around its center at zero-crossings and it ranged from 133 to 333 ms in 20 ms steps. Then the pitch contour of each token was edited (HH<u>H</u>→HL<u>L</u>, LH<u>L</u>, LH<u>HL</u>, LH<u>H</u> (L=low pitch)) by reference to the F0 data (duration of the underlined portions ranged). The four patterns are possible in Tokyo Japanese except a falling tone was realized even when vowel duration was rather short (e.g. 133 ms).

For the F0 data (pitch points and the F0 values), the speaker pronounced /mamama/ and /mamama:/ with three or four possible accent patterns in isolation three times. One token for the four patterns (HLL, LHL, LHH and LHHL) was selected respectively from the three utterances. Three out of four were the tokens having a final short vowel but LHHL, which has a final long vowel, was used because Tokyo Japanese does not allow a short vowel with a contour tone.

The F0 data were extracted using a ManipulationEditor-Stylize pitch (2 st) with Praat. Pitch point (1) and (3) were at the onset of the first and second vowel, pitch point (2), (4) and (6) were at the offset of each vowel. Pitch point (5) was at one pitch period after the onset of the final vowel. Table 3 shows the pitch points and the F0 values.

Table 3: Pitch points and F0 values (Hz).

pitch points	(1)	(2)	(3)	(4)	(5)	(6)
HLL	157.4	173.0	-	-	-	96.8
LHL	131.8	124.0	-	159.0	-	103.5
LHHL	132.0	127.5	152.4	-	157.7	99.7
LHH	131.7	129.0	148.4	-	-	147.0

2.3. Procedure

The participants were tested individually in a quiet room. They identified the final vowel length by clicking the button, “MA” or “MAA” on a computer screen. Each stimulus was presented ten times in random order over loudspeakers, and the test consisted of 440 trials (44 stimuli × 10 times) divided into 22 blocks. The participants took a practice section (16 trials, 2 endpoints) prior to the test section (29 minutes + break).

3. RESULTS AND DISCUSSIONS

Figure 1, Figure 2 and Figure 3 show percentages of responses as a long vowel (% long) exhibited by NJ, NC-higher and NC-lower, respectively. As it is clear from the figures, their % long increased as function of vowel duration. The result indicates that listeners in the three groups attend to durational cue to identify length even though NC do not use duration as a distinctive cue in their L1.

Figure 1: % long by the NJ group.

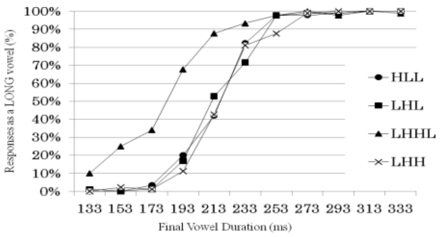


Figure 2: % long by the NC-higher group.

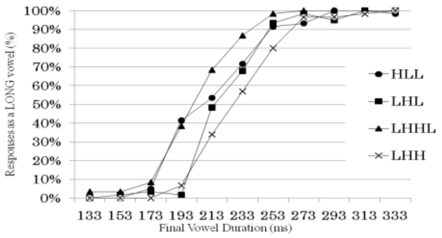
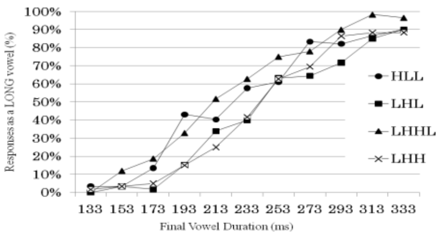


Figure 3: % long by the NC-lower group.



Probit analyses were performed on each listener's identification function to estimate the boundary location (50% crossover point) and the boundary width (75% - 25% crossover points).

3.1. Boundary width

In the figures changes from one category to another were more abrupt for NJ and NC-higher than for NC-lower. NC-lower's % long increased gradually with vowel duration. This may indicate that perception by NC-lower is not categorical but continuous, which partially supports [11].

Table 4 demonstrates the mean boundary widths for each group. Narrower boundary widths indicate sharper judgment, which reflects categorical perception (CP) in one sense. The results of an ANOVA for the NC group (between factor: level, within factor: pitch) showed significant main effects of level and pitch [level: $F(1,10)=12.94$, $p<.01$, pitch: $F(3,20)=3.95$, $p<.05$]. NC-lower's boundary width (58.5) was significantly wider than NC-higher's (32.3) and a post hoc test revealed that the width of HLL (56.3) was significantly wider than that of LHL (37.9) ($p<.05$). An ANOVA for the NJ and NC-lower's data (between factor: L1, within factor: pitch) showed a significant main effect of L1 [$F(1,13)=14.68$, $p<.01$] indicating that NJ's width (30.2) was significantly narrower than NC-lower's (58.5). Neither main effect nor interaction was found on an ANOVA for the data of NJ and NC-higher (between factor: L1, within factor: pitch) and an ANOVA for NJ's data (within factor: pitch) did not show a main effect.

Table 4: Mean boundary widths (ms).

	NJ		NC-higher		NC-lower	
		SD		SD		SD
HLL	28.5	15.0	40.5	13.7	72.1	28.4
LHL	28.0	10.0	27.3	14.4	48.5	15.8
LHHL	37.9	29.1	29.4	9.6	61.2	20.4
LHH	26.3	9.5	31.9	11.5	52.3	17.8
Mean	30.2	15.9	32.3	12.3	58.5	20.6

NC-higher's boundary width was not different from NJ's but NC-lower's was significantly wider than the other two. These results support [4]'s claim and display the capability of approximating perception function in the L2 to native speakers', which is consistent with the claim in [8]. [11] says that CP is the very problem for NC, but the current study shows that it is possible to perceive L2 contrasts categorically when learning progresses.

Pitch did affect NC's boundary width but not NJ's and the width of HLL was significantly wider than that of LHL. [8] showed that NC's % correct of length identification tended to be lower when the first vowel was accented. The wider boundary width can explain the lower % correct, but it is unclear what causes this phenomenon in HLL.

3.2. Boundary location

The curves of % long demonstrated by the NJ and NC-higher group generally look similar in that % long increased from 0 to 100%; however, those of the NC-lower evidently differs in that their % long did not reach 100% even when the duration was long enough for NJ and NC-higher to identify length as a long vowel. This means that NC-lower require longer duration to perceive a long vowel.

Table 5 shows the mean boundary values. The results of an ANOVA for the NC group showed a significant main effect of pitch [$F(3,30)=3.41$, $p<.05$]. A post hoc test revealed that the boundary value of LHHL (210.6) was significantly smaller than that of LHH (236.2). An ANOVA for the NJ and NC-lower's data showed significant main effects of L1 and pitch [L1: $F(1,13)=7.02$, $p<.05$, pitch: $F(3,39)= 7.66$, $p<.001$]. The results indicate that NC-lower's value (237.1) was significantly greater than NJ's (206.5), in addition, the value of LHHL (199.2) was significantly smaller than that of HLL (222.5, $p<.05$), LHL (232.1, $p<.01$) or LHH (233.3, $p<.001$). The results of an ANOVA for the data of NJ and NC-higher demonstrated a significant main effect of pitch [$F(3,39)=29.4$, $p<.001$] and interaction between L1 and pitch [$F(3,39)=3.41$, $p<.05$]; therefore, ANOVAs (within factor: pitch) were conducted further for NJ and NC-higher, separately. The ANOVA for NJ's data showed a main effect of pitch and a post hoc test revealed that the value of LHHL (178.3) was significantly smaller than the others (all $p<.001$). The ANOVA for the data of NC-higher also indicated a main effect of pitch and a post hoc test showed LHHL (201.1) < LHL (221.1, $p<.01$), LHH (229.6, $p<.001$) and HLL (213) < LHH (229.6, $p<.01$).

Table 5: Mean boundary location (ms).

	NJ		NC-higher		NC-lower	
		SD		SD		SD
HLL	214.4	8.6	213.0	17.2	230.7	51.5
LHL	215.3	10.7	221.2	15.7	251.2	46.3
LHHL	178.3	25.4	201.1	15.6	220.1	27.8
LHH	217.8	13.7	229.6	14.1	246.4	38.3
Mean	206.5	14.6	216.2	15.7	237.1	41.0

NC-higher's mean boundary value was not different from NJ's but NC-lower's was significantly and numerically greater than NJ's and NC-higher's, respectively. Therefore, it is possible to think that NC set a greater boundary at an early stage of learning but modify and approximate it to NJ's later. This is inconsistent with [4] and it may be due to the difference in the grouping method.

Pitch affected listeners in the three groups in the same manner and a falling tone increased their perceived vowel duration. It does not seem, however, that Chinese has contributing factors to this increase. [5] says that the influence of a dynamic F0 on the perception appears only in native listeners of languages which associate a dynamic F0 with longer vowel duration. One possibility is that NC have learned to use a contour tone for identification and the distinctive role of pitch in their L1 have helped them learn to use it.

NC's boundary location of HLL tended to be smaller than that of LHH and the difference was significant for NC-higher. [7] and [8] report that final vowel duration is shorter when a word is accented than it is unaccented in NJ's production. Therefore, HLL < LHH on NC's boundary seems to reflect and correlate with such NJ's production.

4. CONCLUSIONS

NC-lower showed wider boundary width and greater boundary location; on the other hand, NJ and NC-higher were not significantly different in their boundary location and width. These results suggest that non-native speakers' perception function in the L2 approximate to native speakers' (i.e. CP) as L2 learning proceeds. Future study should include measurements of reaction time to attest the approximation more precisely. Pitch affected NC's perception in the same way as it did to NJ's. NC seem to have learned such perception; moreover they seem to have modified their perception in accordance with NJ's production.

5. ACKNOWLEDGEMENTS

I thank Mr. Hajime Takeyasu for his technical support on the experiment and the statistic analyses. This study was partly supported by Grant-in-Aid for JSPS Fellows (#22-365).

6. REFERENCES

[1] Boersma, P., Weenink, D. 2009. Praat: doing phonetics by computer (Version 5.1.04), [Computer program]. Online: *http://www. praat.org/*.

[2] Fujisaki, H., Sugito, M. 1977. Onsei no butsuriteki seishitsu. In Ono, S., Shibata, T. (eds.), *Iwanami Koza Nihongo 5 On'in*. Tokyo: Iwanamishoten, 63-98.

[3] Kinoshita, K., Behne, D.M., Arai, T. 2002. Duration and F0 as perceptual cues to Japanese vowel quantity. *Proc. 7th ICSLP* Denver.

[4] Kurihara, M. 2004. Perceptual characteristics of Japanese long vowels by Mandarin Chinese speakers. *Journal of Linguistic Science (Tohoku University)* 8, 1-12.

[5] Lehnert-LeHouillier, H. 2007. The influence of dynamic F0 on the perception of vowel duration: Cross-linguistic evidence. *Proc. 16th ICPhS* Saarbrücken, 757-760.

[6] Minagawa, K.Y., Maekawa, K., Kiritani, S. 2002. Effects of pitch accent and syllable position in identifying Japanese long and short vowels: Comparison of English and Korean speakers. *J. Phone. Soc. Jap.* 6(2), 88-97.

[7] Mori, Y. 2001. Effect of accentual fall on final lengthening in Japanese. *J. Phone. Soc. Jap. 5*(1), 92-106.

[8] Takiguchi, I. 2010. *Second Language Speech Perception: Pitch Cues to Japanese Vowel Length.* Unpublished MA thesis, Kobe University.

[9] Takiguchi, I., Takeyasu, H., Giriko, M. 2010. Effects of a dynamic F0 on the perceived vowel duration in Japanese. *Proc. 5th Speech Prosody* Chicago, 100944.

[10] Toda, T. 2003. *Second Language Speech Perception & Production: Acquisition of Phonological Contrasts in Japanese.* Lanham. Maryland: Univ. Press of America.

[11] Uchida, T. 1993. Characteristics of auditory cognition of long vowels and double consonants for Chinese students in learning Japanese language. *Japanese Journal of Educational Psychology*, 41, 414-423.