Perception of Japanese quantity by Swedish speaking learners: A preliminary analysis

Miyoko Inoue
Department of Japanese Language & Culture, Nagoya University, Japan

Abstract

Swedish learners' perception of Japanese quantity was investigated by means of an identification task. Swedish informants performed similarly to native Japanese listeners in short/long identification of both vowel and consonant. The Swedish and Japanese listeners reacted similarly both to the durational variation and to the F0 change despite the different use of F0 fall in relation with quantity in their L1.

Introduction

A quantity language is a language that has a phonological length contrast in vowels and/or consonants. Japanese and Swedish are known as such languages, and they employ both vowels and consonants for the long/short contrast (Han, 1965, for Japanese; Elert, 1964, for Swedish). Both of the languages use duration as a primary acoustic cue to distinguish the long/short contrast.

Quantity in Japanese is known to be difficult to acquire for learners (e.g. Toda, 2003), however, informants in previous research have mainly been speakers of non-quantity languages. In their research on L2 quantity in Swedish, McAllister et al. (2002) concluded that the degree of success in learning L2 Swedish quantity seemed to be related to the role of the duration feature in learners' L1. It can, then, be anticipated that Swedish learners of Japanese may be relatively successful in acquiring Japanese quantity. The present study, thus, aims to investigate whether Swedish learners are able to perform similarly to the Japanese in the perception of Japanese quantity of vowels and consonants.

In addition to duration, there can be other phonetic features that might supplement the quantity distinction in quantity languages. In Swedish, quality is such an example, but such a feature may not be necessarily utilized in other languages. For example, quality does not seem to be used in the quantity contrast in Japanese (Arai et al., 1999).

Fundamental frequency (F0) could be such a supplementary feature in Japanese. Kinoshita et al. (2002) and Nagano-Madsen (1992) reported that the perception of quantity in L1 Japanese was affected by the F0 pattern. In their experiments, when there was a F0 fall within a vowel, Japanese speakers tended to perceive the vowel as 'long'. On the other hand, a vowel with long duration was heard as 'short' when the onset of F0 fall was at the end of the vowel (Nagano-Madsen, 1992).

These results are in line with phonological and phonetic characteristics of word accent in Japanese. It is the first mora that can be accented in a long vowel, and the F0 fall is timed with the boundary of the accented and the post-accent morae of the vowel. Since the second mora in a long vowel does not receive the word accent, a F0 fall should not occur at the end of a long vowel.

In Swedish, quantity and word accent seem only to be indirectly related to the stress in such a way that the stress is signaled by quantity and the F0 contour of word accent is timed with the stress. In the current research, it is also examined if Swedish learners react differently to stimuli with and without F0 change. Response to unaccented and accented words will be compared. An unaccented word in Japanese typically has only a gradual F0 declination, while an accented word is characterized by a clear F0 fall. It can be anticipated that Swedish learners would perform differently from native Japanese speakers.

Methodology

An identification task was conducted in order to examine the categorical boundary between long and short vowels and consonants and the consistency of the categorization. The task was carried out in a form of forced-choice test, and the results were compared between Swedish and Japanese informants.

Stimuli

The measured data of the prepared stimuli[1] are shown in Table 1 and

Table 2. The original sound, accented and unaccented /mamama, mama:ma/ (for long/short vowel) and /papapa, papap:a/ (for long/short consonant), was recorded by a female native Japanese speaker. For the accented version, the 2nd mora was accented for both *ma-* and *pa*-series. The stimuli were made by manipulating a part of recorded tokens with Praat (Boersma and Weenink, 2004) so that the long sound shifts to short in 7 steps. Thus, a total of 28 tokens (2 series x 7 steps x 2 accent type) were prepared. The F0 peak, the location of F0 peak in V2 and the final F0 were fixed at the average value of long and short sounds.

Table 1. The measurements of the stimuli in ma-*series (adopted from Kanamura, 2008: 30 (Table 2-2) and 41 (Table 2-5), with permission). The unaccented and the accented stimuli are differentiated by the utterance final F0 (rightmost column).*

No.	Ratio	C3 Duration (ms)	Word Duration (ms)	F0 Peak (Hz)	Peak Location in V2	Final F0 (Hz)
1	0.25	78	582			
2	0.40	128	627			
3	0.55	168	673			242 (unacc)
4	0.70	213	718	330	48%	
5	0.85	259	764			136 (acc)
6	1.00	303	810			
7	1.15	349	855			

Table 2. The measurements of the stimuli in pa-*series. The unaccented and the accented stimuli are differentiated by the utterance final F0 (rightmost column).*

No.	Ratio	C3 Duration (ms)	Word Duration (ms)	F0 Peak (Hz)	Peak Location in V2	Final F0 (Hz)
1	0.25	85	463			
2	0.40	136	514			
3	0.55	188	566			231 (unacc)
4	0.70	239	617	295	96%	
5	0.85	290	668			116 (acc)
6	1.00	341	719			
7	1.15	392	770			

Informants

The informants were 23 Swedish learners of Japanese (SJ) at different institutions in Japan and Sweden. The length of studying Japanese varied from 3 to 48 months.[2] Thirteen native speakers of standard Japanese (NJ) also participated in the task in order for comparison.

Procedure

An identification task was conducted using ExperimentMFC of Praat. Four sessions (*ma-*/*pa*-series x 2 accent) were held for each informant. In each session, an informant listened to 70 stimuli (7 steps x 10 times) randomly and answered whether the stimulus played was, for example, /mamama/ or /mama:ma/ by clicking on a designated key.

Calculation of the categorical boundary and the 'steepness' of categorical function

The location of the categorical boundary between long and short, and also the consistency ('steepness') of the categorization function was calculated following Ylinen et al. (2005). The categorical boundary is indicated in milliseconds. The value of steepness is interpreted in such a way that the smaller the value, the stronger the consistency of the categorization function.

Results

It was reported in Kanamura (2008) that several of the Chinese informants did not show correspondence between the long/short responses and the duration of V2 in the mamama/mama:ma stimuli. She excluded the data of such informants from the analysis. No such inconsistency between the response and the duration was found for the Swedes or for the Japanese in the current study, and thus none of the data were omitted in this regard. However, the data of one Japanese informant was eliminated from the result of *ma*-series since the calculated boundary location of that person was determined as extremely high compared with the others.

Perception of long and short vowel (*ma*-series)

Figure 1 indicates the percentage of 'short' response to the stimuli. The leftmost stimulus on x-axis (labeled as '0.25') is the shortest sound and the rightmost the longest ('1.15'). The plotted responses traced s-shaped curves, and the curves turned out to be fairly close to each other. Differences are found at 0.55 and 0.70, and the 'short' responses by NJ differ visibly between unaccented and accented stimuli. The 'short' response to the accented stimuli at 0.55

dropped to a little below 80%, but the unaccented stimuli remained as high as almost 100%. SJ's responses to unaccented/accented stimuli at 0.55 appeared closely to each other. Besides, the s-curves of SJ looked somewhat more gradual than those of NJ.

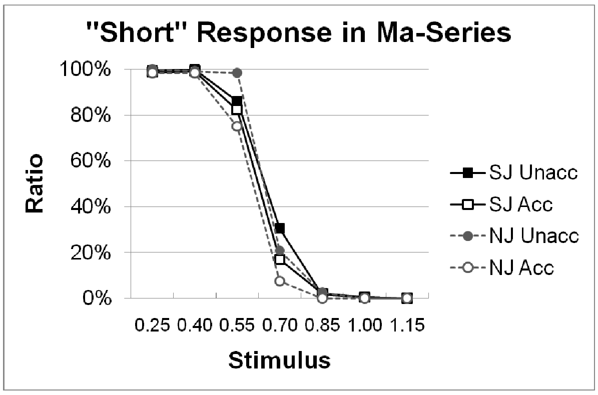

Figure 1. The percentage of "short" responses for stimuli with the shortest to the longest V2 (from left to right on x-axis).

Table 3 shows the mean category boundary and the steepness of the categorization function. Two-way ANOVAs were conducted for the factors Group (SJ, NJ) and Accent Type (Unaccented, Accented) separately for the category boundary and the steepness.

Table 3. The category boundary location (ms) and the steepness of the categorization function in the unaccented (flat) and the accented stimuli of ma*-series.*

	Unaccented (flat)		Accented	
	SJ	NJ	SJ	NJ
Boundary (ms)	199.6	200.0	191.3	182.6
SD	13.6	15.1	13.4	13.2
Steepness	27.8	16.3	27.6	18.8
SD	7.7	8.9	9.7	10.5

For the categorical boundary, the interaction between Group and Accent Type tended to be significant ($F(1,33)=3.31$, $p<.10$). There was a simple main effect of Group at Accented condition which was also tended to be significant ($F(1,33)=3.15$, $p<.10$). Meanwhile, simple main effects of Accent Type at SJ ($F(1,33)=5.57$, $p<.05$) and at NJ ($F(1,33)=24.34$, $p<.01$) were significant. For both of the speaker groups, the location of the perceptual boundary was earlier in accented than unaccented words. In sum, there was no difference between the two groups, but there was between the two accent types. 'Long' responses were invoked by shorter duration of V2 of the accented stimuli.

As for the steepness, there was no significant interaction for the factors Group and Accent Type. A significant main effect was found for Group ($F(1,34)=11.48$, $p<.01$). The mean of steepness was greater for SJ than for NJ. This means that the categorization function of NJ is more consistent than that of SJ as in the above-interpretation of Figure 1.

Perception of long and short consonant (*pa*-series)

The ratio of 'short' response to the stimuli in *pa*-series is given in Figure 2. As in *ma*-series, the plotted responses make s-shaped curves. Noticeable difference in the 'short' response rate was found at 0.55 and 0.70 on x-axis. The difference between SJ and NJ seems to be greater for the accented stimuli, and SJ's 'short' response drops below 60% at 0.55 while NJ marks around 80%, which makes SJ's curve more gradual than NJ's. But the curves of SJ and NJ for Unaccented stimuli almost overlap with each other.

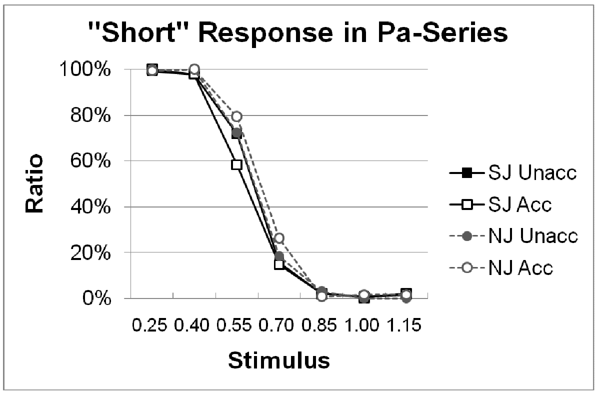

Figure 2. The percentage of "short" responses for stimuli with the shortest to the longest C3 (from left to right on x-axis).

The categorical boundary and the steepness are shown in Table 4. Two-way mixed ANOVAs were carried out for the categorical boundary and the steepness with the factors Group (SJ, NJ) and Accent Type (Unaccented, Accented). The categorical boundary exhibited a tendency of interaction between the factors ($F(1,34)=3.50$, $p<.10$). A simple main effect of Group at Accented condition was also found to tend to be significant ($F(1,34)=3.67$, $p<.10$). It was natural that difference caused by Accent Type did not reach statistical significance because the consonant duration here corresponded to the closure (silence) duration, and there was of course no F0 information relevant to the quantity.

The result for steepness was similar to that of *ma*-series. There was no significant interaction between the factors Group and Accent Type (F(1,34)=0.00, *n.s.*), but there was a significant main effect of Group (F(1,34)=11.47, *p*<.01). The mean of NJ was smaller than that of SJ. Thus, the Japanese made a clearer distinction between long and short than the Swedes.

Table 4 The category boundary location (ms) and the steepness of the categorization function in the unaccented (flat) and the accented stimuli of pa-*series.*

	Unaccented (flat)		Accented	
	SJ	NJ	SJ	NJ
Boundary (ms)	210.2	210.0	203.7	219.5
SD	22.0	31.0	19.6	29.8
Steepness	36.2	22.0	39.9	26.1
SD	15.5	8.7	16.3	21.1

Conclusion

The Swedish informants generally performed similarly to the Japanese. They were quite successful as opposed to what has been reported on other L1 speakers. This might be due to the use of duration in L1 phonology in Swedish, but further inter-linguistic comparison is necessary to obtain a clearer view.

It was expected that the two speaker groups would perform differently for the accented stimuli, but this was not true. The F0 change caused the earlier long/short boundary for both of the groups. Further investigation is needed to find out if this has been acquired during the course of learning or if it originated from the characteristics in Swedish.

Acknowledgements

I would like to thank Kumi Kanamura for the permission to use the sound stimuli in her research in Kanamura (2008), and also the students and teachers at University of Gothenburg, Lund University/Gifu University, Nagoya University and Stockholm University for their kind cooperation in the experiments. The research presented here was supported by The Scandinavia-Japan Sasakawa Foundation for the year 2008.

Notes

1. The stimuli of ma-*series were adopted from Kanamura (2008) with permission.*

2. According to personal interviews, Japanese classes at high schools were usually held only once a week as opposed to several hours per week at universities. The length of studying Japanese excluding high school is 3-24 months.

References

Arai, T., Behne, D. M., Czigler, P. and Sullivan, K. (1999) Perceptual cues to vowel quantity: Evidence from Swedish and Japanese. Proceedings of the Swedish Phonetics Conference. Fonetik 99 (Göteborg) 81, 8-11.

Boersma, P. and Weenink, D. (2004) Praat: Doing phonetics by computer (Version 4.2) [Computer program]. Retrieved March 4, 2004, from http://www.praat.org.

Elert, C.-C. (1964) Phonologic Studies of Quantity in Swedish. Stockholm: Almqvist & Wiksell.

Han, M. S. (1965) The feature of duration in Japanese. Study of Sound (Journal of Phonetic Society of Japan) 10, 65-80.

Kanamura, K. (2008) Why the acquisition of long and short vowels in L2 Japanese is difficult? : The perception and recognition of Japanese long and short vowels by Chinese learners [In Japanese]. Doctoral dissertation. Nagoya University.

Kinoshita, K., Behne, D. M. and Arai, T. (2002) Duration and F0 as perceptual cues to Japanese vowel quantity. Proceedings of the International Conference on Spoken Language Processing. ICSLP-2002 (Denver, Colorado, USA), 757-760.

McAllister, R., Flege, J. E. and Piske, T. (2002) The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. Journal of Phonetics 30, 229-258.

Nagano-Madsen, Y. (1992) Mora and prosodic coordination: A phonetic study of Japanese, Eskimo and Yoruba. Travaux de l'Institut de Linguistique de Lund 27. Lund: Lund University Press.

Toda, T. (2003) Second Language Speech Perception and Production: Acquisition of Phonological Contrasts in Japanese. Lanham, MD: University Press of America.

Ylinen, S., Shestakova, S., Alku, P. and Huotilainen, M. (2005) The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. Language and Speech 48 (3), 313-338.