SENSITIVITY TO PALATALIZED SEQUENCES DOES NOT TRANSFER TO NON-NATIVE PALATALIZED CONTRASTS

Risa Matsubara, Izabelle Grenon

The University of Tokyo
matsubara-risa761@g.ecc.u-tokyo.ac.jp, grenon@boz.c.u-tokyo.ac.jp

ABSTRACT

This study investigates the acquisition of the /CʲV/-/CʲjV/-/CʲijV/ palatalized contrasts in Russian by sixteen Japanese speakers of low or high level of proficiency. The Russian palatalized sequences contrast from each other by the duration of the F2 steady state, a cue used by Japanese speakers to distinguish /CʲV/ from /CiV/ in their native language. Considering their sensitivity to this cue and their longer Russian experience, high proficiency learners were expected to produce the novel contrast /CʲjV/-/CʲijV/ more accurately than the low proficiency group. The results revealed that all learners could make a distinction between /CʲV/ and /CʲjV/, possibly by producing the latter sequence as /CijV/ or /CʲijV/ to make it conform to Japanese phonotactics, while neither group could significantly contrast /CʲjV/ from /CʲijV/. Hence, prior sensitivity to the duration of the F2 steady state in their native language did not facilitate the acquisition of non-native palatalized contrasts by Japanese speakers.

Keywords: Speech production, palatalized consonants, Russian, Japanese.

1. INTRODUCTION

Russian features a three-way palatalized contrast, with /CʲV/ as in /su**dʲa**/ ‘judging’ (adverbial participle, present), /CʲjV/ as in /su**dʲja**/ ‘judge’, and /CʲijV/ as in /su**dʲija**/ ‘judge’ (archaic). These sequences are acoustically contrasted from each other by the length of the F2 steady state before the formant transition, with /CʲjV/ having a longer steady state than /CʲV/ [10], and /CʲijV/ exhibiting a longer steady state than /CʲjV/ [4]. Palatalized sequences may be challenging for second language (L2) learners to contrast if their first language (L1) lacks palatalized segments (e.g., [4]). Conversely, learners who can use the length of the F2 steady state contrastively in their L1 may show a facilitation effect for the acquisition of the Russian palatalized contrasts. The current study evaluates this hypothesis with Japanese speakers classified into two levels of proficiency in Russian (low and high). All learners were expected to contrast the palatalized contrast /CʲV/-/CʲjV/, possibly by replacing the illicit /Cʲj/ sequence by /Cij/ or /CʲijV/. Provided that L2 experience may have an effect on the acoustic realization of L2 contrasts [5], and that they should already be sensitive to the critical acoustic cue that distinguishes the Russian palatalized sequences, the high proficiency learners were expected to produce a distinction between /CʲjV/ and /CʲijV/.

Japanese features palatalized /CʲV/ sequences [1, 8], such as /kʲu:/ ‘immediate’, which contrast with /CʲijV/ sequences, such as /kʲiju:/ ‘groundless fears’. Since this contrast exists in Japanese, the production of the /CʲV/-/CʲijV/ contrast in Russian was not analyzed in the current study. In addition, a sequence such as /CʲV/ in Japanese is contrasted from the sequence /CiV/ by the length of the F2 steady state before the formant transition, where a steady state of less than 50ms is generally perceived as the presence of a glide [11]. Hence, Japanese speakers are sensitive to the difference of the length of the F2 steady state to make contrast in their L1, which may facilitate the acquisition of non-native palatalized contrasts.

On the other hand, Japanese phonotactics do not allow consonant clusters, such as /CC/ or /CʲC/. In order to avoid these clusters, Japanese speakers are known to use vowel epenthesis in loanword adaptation [6], where a word like ‘grapefruit’ /grepfrut/ in English is usually adapted as /g**u**re:p**u**f**u**ru:tsu/. When /Cʲ/ is included in the cluster, the palatalization may be replaced by a full vowel /i/, resulting in the sequence /Ci/ instead. This adaptation is observed in Russian loanword in Japanese, where the word /arxangelʲsk/ ‘Arkhangelsk (a city in Russia)’ is adapted as /aruhanger**i**suku/ [9]. Hence, the Russian sequence /CʲjV/ may be produced by the L2 learners as either /CijV/ or /CʲijV/ if they cannot produce the sequence as /CʲjV/.

Accordingly, our first hypothesis was that the /CʲV/-/CʲjV/ contrast should be produced contrastively by all learners, whether the sequence /CʲjV/ is produced properly, with an epenthetic vowel or with the palatalization replaced by a full vowel (in the latter two cases the F2 steady state would be lengthen and the length of the vocalic interval would be close to that of the /CʲijV/ sequence). Our second hypothesis was that low proficiency learners would be unable to contrast the /CʲjV/-/CʲijV/ sequences as they are expected to adapt the /CʲjV/ sequence to

Japanese phonotactics. Conversely, the high proficiency learners were expected to make some contrast in the duration of the glide+vowel interval between /C^{j}jV/ and /C^{j}ijV/, provided that prior sensitivity to the length of the F2 steady state in their L1 facilitates the acquisition of L2 contrasts.

2. METHOD

2.1. Participants

Nine native Russian speakers aged 24-34 (five males, mean: 26.9, SD: 3.3), and 16 native Japanese speakers learning Russian aged 19-23 (mean: 21.4, SD: 1.1), took part in the experiment. None of the participants reported any history of speech or hearing impairment. A monetary compensation was given for their participation.

The Japanese participants were divided into two proficiency groups: 7 were assigned to the low proficiency group (six males), and 9 to the high proficiency group (one male). Their proficiency level was determined by their length of Russian education. The participants assigned to the low proficiency group had no more than 2 years of Russian education (mean: 1.4 year, SD: 0.6), while the participants assigned to the high proficiency group had at least 2.5 years of Russian education (mean: 3.9 years, SD: 1.3). Six high proficiency learners had experienced studying in a Russian-speaking region (mean: 5.6 months, SD: 4.8).

2.2. Stimuli

The stimuli for the production experiment consisted of eight target nonsense words, shown in Table 1, mixed with eleven non-palatalized filler words. The /CV/, /C^{j}V/, and /C^{j}jV/ words should be stressed on the first syllable, while the /C^{j}ijV/ words should be stressed on the second syllable to prevent vowel reduction of an unstressed vowel (in Russian, unstressed /^{j}a/ and /ja/ are pronounced as [i] or [ɪ] [8]). Each word was presented in the carrier sentence /ona skazala ___ op^{j}at^{j}/ meaning 'she said ___ again', with the target word underlined, and stress overtly marked except for /C^{j}o/, /C^{j}jo/, and /C^{j}ijo/ words, since /^{j}o/ and /jo/ are obligatorily stressed [8]. The words were presented to participants in the Russian orthography, which distinguishes the palatalization contrasts (e.g., <u>бя́да</u> for /b^{j}ada/, <u>бья́да</u> for /b^{j}jada/, <u>бия́да</u> for /b^{j}ijada/, and <u>бёда</u> for /b^{j}oda/).

The voiced plosive /b/ was selected for the onset consonant because bilabial consonants do not include any other movements than palatal glide in co-articulation and therefore the behavior of the palatal glide is easier to observe [10]. The vowels /a/ and /o/ were chosen because they are common in both Japanese and Russian. Note that only the /C^{j}V/-/C^{j}jV/ and /C^{j}jV/-/C^{j}ijV/ contrasts were analyzed, as we were interested in whether the Japanese speakers could make any duration difference between these two contrasts, as justified in the introduction.

Table 1: /CV/ and palatalized nonsense words used for the experiment.

CV	**C^{j}V**	**C^{j}jV**	**C^{j}ijV**
bada	b^{j}ada	b^{j}jada	b^{j}ijada
boda	b^{j}oda	b^{j}joda	b^{j}ijoda

2.3. Procedure

The participants were recorded with a Shure SM10A low-impedance, unidirectional dynamic microphone at 44,100Hz, in a sound-proof booth at The University of Tokyo, directly to Macbook Pro computer using Praat [3]. The nineteen words used for the experiment were embedded in the carrier sentence and shown to participants in a pseudo-random order one at a time using Microsoft Office PowerPoint. The participants were asked to read all 19 sentences contained in a block, and there were 5 blocks in total (i.e., the 19 sentences were presented 5 times). The speaker had to press a key on the computer keyboard to move on to the next slide (sentence) within a block, but a researcher was in charge to start each new block. Between blocks, the speaker could take a short break. A total of 750 target samples were produced by the speakers, of which 28 were rejected because they were mispronounced (25 speakers x 6 target words x 5 trials = 750 – 28 misread tokens = 722 tokens for analyses).

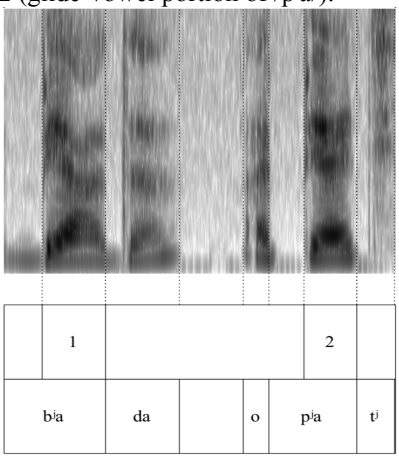
2.4. Analysis

Dividing the palatalized sequence into the F2 steady state and transition may be unreliable in some cases. Assuming that the duration of a syllable is proportional to the duration of the palatal glide included in the syllable [12, 14], we decided to measure the intervals from the onset of the stop release to the end of the vowel of the target syllable (i.e., excluding the stop closure). For instance, the region measured in the sequence /b^{j}a/ corresponds to region 1 in Figure 1, which shows a spectrogram of the sequence /b^{j}ada op^{j}at^{j}/ produced by a Russian speaker. In order to cancel the possible effect of speech rate, the duration of region 1 was divided by the duration of region 2, which is equivalent to /p^{j}a/ (without the stop closure) in the last word /op^{j}at^{j}/. The measurements were done in Praat [3]. The resulting ratios were used in the statistical analyses.

LME (linear mixed-effect model) was used for the statistical analyses. A two by two model was adopted, with Glide Type and Group (low vs. high; low vs.

native; high vs. native) as factors, and the ratios as dependent values. The random factors were participants and items. Backward selection [2] was used for the best-fit model selection.

Figure 1: Spectrogram of the sequence /bʲada opʲatʲ/ produced by a female Russian speaker. Ratio were region 1 (glide-vowel portion of /bʲa/) divided by region 2 (glide-vowel portion of /pʲa/).



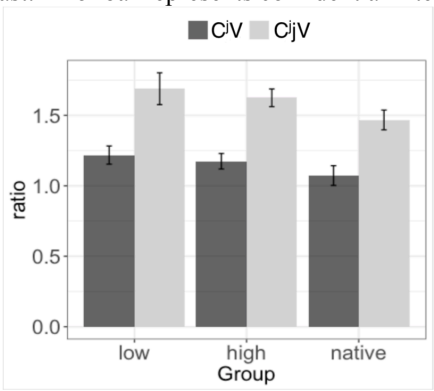
3. RESULTS

As shown in Figure 2, the ratio of /CʲjV/ is larger than that of /CʲV/ in all groups, meaning the interval measured was produced longer in the /CʲjV/ sequences. The LME models with Glide Type (/CʲV/-/CʲjV/) and Group found a significant main effect of Glide Type in all of the analyses reported in Table 2. No significant Glide Type x Group interaction were found in any of the group comparisons, suggesting no difference in performance between the groups. Simple main effect tests were further conducted and revealed that the effect of Glide Type (/CʲV/-/CʲjV/) was significant in all groups (low: $p < 0.05$; high: $p < 0.001$; native: $p < 0.001$), suggesting that all groups were able to make a contrast between these sequences.

Table 2: LME results /CʲV/ vs. /CʲjV/ (Glide Type).

	β	SE	df	*t*	*p*	
	Group = low vs high					
(Intercept)	1.415636	0.048434	15.9415	29.228	2.83E-15	***
Glide Type	0.447672	0.058636	15.62648	7.635	1.18E-06	***
Group	0.030104	0.097642	15.96891	0.308	0.762	
Glide Type:Group	0.004063	0.122563	15.09956	0.033	0.974	
	Group = low vs native					
(Intercept)	1.33793	0.05518	15.93886	24.246	5.24E-14	***
Glide Type	0.42752	0.05752	15.5569	7.432	1.69E-06	***
Group	-0.16371	0.11124	15.95913	-1.472	0.161	
Glide Type:Group	-0.0413	0.11986	14.94058	-0.345	0.735	
	Group = high vs native					
(Intercept)	1.33545	0.04325	8.39417	30.874	6.12E-10	***
Glide Type	0.42733	0.04775	17.81377	8.949	5.22E-08	***
Group	-0.13354	0.07395	17.71162	-1.806	0.088	.
Glide Type:Group	-0.03957	0.0955	17.81706	-0.414	0.684	

Figure 2: Difference in ratio between the low (left), and high proficiency learners (middle), and native Russian speakers (right) on the /CʲV/ and /CʲjV/ contrast. Error bar represents confidential interval.

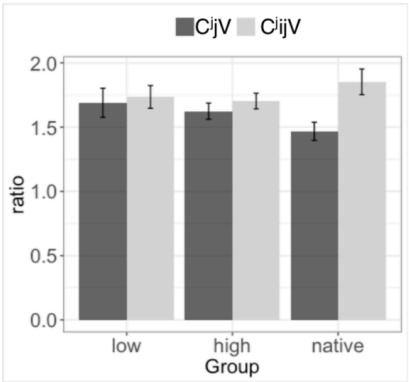


Conversely, as seen in Figure 3, neither the low nor the high proficiency group was able to contrast /CʲjV/ from /CʲijV/. The Japanese groups performed similarly to each other, but differently from native Russian speakers.

Table 3: LME results /CʲjV/ vs. /CʲijV/ (Glide Type).

	β	SE	df	*t*	*p*	
	Group = low vs high					
(Intercept)	1.67216	0.07012	11.61435	23.846	3.16E-11	***
Glide Type	0.07256	0.04394	14.53082	1.651	0.12	
Group	0.02122	0.14933	9.14027	0.142	0.89	
Glide Type:Group	-0.01504	0.09165	13.53696	-0.164	0.872	
	Group = low vs native					
(Intercept)	1.67121	0.07475	15.9819	22.357	1.75E-13	***
Glide Type	0.24496	0.04612	13.75852	5.311	0.000117	***
Group	-0.02418	0.15062	16.00384	-0.161	0.87446	
Glide Type:Group	0.3097	0.09584	12.6024	3.231	0.0068	**
	Group = high vs native					
(Intercept)	1.66E+00	5.81E-02	1.80E+01	28.624	< 2e-16	***
Glide Type	2.27E-01	2.79E-02	3.38E+02	8.126	8.46E-15	***
Group	3.53E-05	1.16E-01	1.80E+01	0	1	
Glide Type:Group	2.96E-01	5.58E-02	3.38E+02	5.303	2.06E-07	***

Figure 3: Difference in ratio between the low (left) and high proficiency learners (middle), and native Russian speakers (right) on the /CʲjV/ and /CʲijV/ contrast. Error bar represents confidential interval.



The LME models with Glide Type (/CʲjV/-/CʲijV/) and Group found a significant Glide Type x Group interaction effect and a significant main effect of

Glide Type in the Group comparison with native speakers, as reported in Table 3. Simple main effect test conducted with regards to the significant interaction effect indicates that the main effect of Glide Type is significant only in native speakers ($p < 0.001$). Additionally, no significant interaction effect was found in the comparison of the low proficiency and high proficiency learners ($p = 0.872$).

4. DISCUSSION

It was hypothesized that both low and high proficiency Japanese learners of Russian would be able to produce a significant contrast in glide-vowel duration between the sequence /C^{j}V/—which exists in their L1—and the sequence /C^{j}jV/—which do not exist in their L1 and may accordingly be realized as /C^{j}ijV/ or /CijV/ to fit Japanese phonotactic constraints (and hence, may exhibit a longer glide+vowel duration than /C^{j}V/). The significant main effects of Glide Type in group comparisons and in each group separately support this hypothesis and suggest that both low and high proficiency learners could make a duration contrast between these two palatalized sequences.

It was further hypothesized that a difference would be observed between the low proficiency and high proficiency group on the production of the /C^{j}jV/-/C^{j}ijV/ contrast based on their length of Russian experience. Japanese speakers with more Russian experience—the high proficiency group—were expected to make a significantly larger contrast in glide+vowel duration than the low proficiency group. However, the results do not concur with this hypothesis: The lack of significant effect of Glide Type in each group of learners suggests that neither the low proficiency nor the high proficiency learners could differentiate the /C^{j}jV/- /C^{j}ijV/ contrast in production. Furthermore, the significant effect of Glide Type x Group interaction in low vs. native and high vs. native comparisons indicates that both groups of learners performed differently from native Russian speakers.

The fact that both low proficiency and high proficiency learners could produce a difference between the sequences /C^{j}V/ and /C^{j}jV/ but not between /C^{j}jV/ and /C^{j}ijV/ suggests no effect of language experience on the production of the Russian palatalized sequences by Japanese speakers in the current tested conditions. And this, despite the fact that Japanese listeners are presumably sensitive to the critical acoustic cue that serve to make the palatalized contrasts (the duration of the F2 steady state, which should impact the duration of the glide+vowel interval, measured as region 1 in our experiment). However, the present study was based only on the measurement of the duration of the glide+vowel interval. Detailed analyses looking at the length of the F2 steady state and shape of the F2 transition and other possible correlates (e.g., pitch) may provide different results. In particular, it could confirm whether the sequence /C^{j}jV/ was realized by the learners with vowel epenthesis or with the replacement of palatalization with a full vowel.

Nevertheless, the current results concur with some previous studies. For instance, language experience was found to have no effect on the production of English vowels by adult Japanese speakers after one year of immersion in an English-speaking country [13]. While the current study suggests, furthermore, that the amount of formal language education has little effect on the production of L2 contrasts, intensive phonetic training appears to provide different results. For instance, about half of the Japanese young adults trained with the English high front vowel contrast were able to contrast the vowels after only two to five hours of phonetic training [7]. These results contrast with the results of Oh et al. [13] and the current results, where the L2 learners presumably have been exposed to a considerably larger amount of hours in the L2. Hence, a focus on the acoustic forms may be required in order to change the perception and production of L2 contrasts, even when the learners are sensitive, to some extent, to the critical acoustic cue that serves to contrast the target non-native sequences.

5. CONCLUSION

This study looked at the effect of formal L2 education on the production of Russian palatalized sequences by native Japanese speakers divided into low proficiency and high proficiency group based on their length of Russian education. It was hypothesized that both groups should be able to contrast the /C^{j}V/ and /C^{j}jV/ sequences, while the high proficiency group should produce the /C^{j}jV/ and /C^{j}ijV/ contrast more accurately than the low proficiency group. Unlike the study by Oh et al. [13], which found no effect of language experience for adult learners, an effect of language experience was expected in our study based on the assumption that Japanese speakers should already be sensitive to the duration of the F2 steady state, which serve to contrast the Russian palatalized sequences. While the /C^{j}V/ and /C^{j}jV/ contrast was produced contrastively by both groups of learners, the /C^{j}jV/ and /C^{j}ijV/ contrast was not contrasted by either group of learners. These results suggest no effect of length of language education on the production of a non-native contrast, even when the learners are presumably sensitive to the cue that serves to contrast the novel segments.

6. REFERENCES

[1] Akamatsu, T. 1997. *Japanese Phonetics: Theory and Practice.* Newcastle: Lincom Europa.

[2] Bates, D., Kliegl, R., Vasishth, S., Baayen, H. 2015. Parsimonious mixed models. arXiv preprint arXiv:1506.04967.

[3] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer. (version 6.0.39) [Computer Programme]. http://www.praat.org/

[4] Diehm, E, E. 1998. Gestures and linguistic function in learning Russian: production and perception studies of Russian palatalized consonants. *Doctoral dissertation, Ohio State University.*

[5] Flege, J. E., Bohn, O.-S., Jang, S. 1997. Effects of experience on non-native speaker's production and perception of English vowels. *J. Phonetics, 25,* 437-470.

[6] Goulet, M.-J. 2001. L'adaptation phonologique d'emprunts lexicaux anglais en japonais et la faible influence de la graphie. *MA Thesis. Laval University, Canada.*

[7] Grenon, I., Kubota, M., Sheppard, C. 2019. The creation of a new vowel category by adult learners after adaptive phonetic training. *J. Phonetics, 72*, 17-34.

[8] Kamiyama, T. 2012. *Roshiago-Onsei-Gaisetsu* [Outline of Russian Phonetics]. Tokyo: Kenkyusha

[9] Kochetov, A. 2008. Phonology and phonetics of loanword adaptation: Russian place names in Japanese and Korean. *Toronto Working Papers in Linguistics, 28*.

[10] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages.* Oxford, UK: Blackwell

[11] Nakata, K., Suzuki, J. 1959. Synthesis and perception of Japanese vowels and vowel-like sounds. *Quarterly review of the Radio Research Laboratories, 5* (21), 237-244.

[12] Nogita, A. 2016. Arguments that Japanese [Cj]s are complex onsets: durations of Japanese [Cj]s and Russian [Cʲ]s and blocking of Japanese vowel devoicing. *Working papers of the Linguistics Circle of the University of Victoria, 26* (1), 73-99

[13] Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., Yamada, T. 2011. A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *J. Phonetics, 39,* 156-167.

[14] Suh, Y., Hwang, J. 2016. The Korean Prevocalic Palatal Glide: A Comparison with the Russian Glide and Palatalization. *Phonetica, 73*(2), 85-100.