Arguments that Japanese [Cj]s are complex onsets: durations of Japanese [Cj]s and Russian [C^{j}]s and blocking of Japanese vowel devoicing

Akitsugu Nogita
University of Victoria
akitsugu@uvic.ca

This study examines whether the Japanese consonant-glide sequence [Cj]s as in [mjakɯ] 'pulse' are complex onsets /CC/ or palatalized consonants /C^{j}/ on the basis of duration. I compared the controversial Japanese consonant-cluster [Cj] analysis and uncontroversial Russian palatalized /C^{j}/ analysis in duration. The results indicate that Japanese [CjV]s are significantly longer than their [CV] counterparts, whereas Russian [C^{j}V]s are not significantly longer than their [CV] counterparts. Thus, in terms of duration, Japanese [Cj]s resemble consonant clusters /CC/ or /Cj/, and not Russian /C^{j}V/s. On the other hand, no arguments seem to exist to support that Japanese [Cj]s are palatalized single consonants based on the results. In addition to duration differences, I also assumed that if [Cj] is a consonant cluster, [j] in [CjɯC] would block [ɯ] devoicing/deletion. The results indicate that [j] blocks [ɯ] devoicing to some extent, but this may also be because infrequent morae may be less frequently devoiced. Therefore, the devoicing pattern does not support the Japanese complex onset hypothesis as strongly as the duration patterns.
Keywords: Japanese consonant-glide sequence; Russian palatalized consonant; complex onset; high vowel devoicing

Introduction

There is a long-standing debate on whether [Cj] sequences, as in [ɾja.kɯ][1] 'omission' in Standard Japanese (SJ), are complex onsets /CC/ (/Cj/) or single palatalized consonants /C^{j}/. In this paper, I attempt to settle this long-standing debate on the basis of phonetic duration. In the first part of this paper, I examine 1) whether SJ [CjV] sequences are phonetically longer than their [CV] counterparts, 2) and if they are, whether the durational difference between SJ [CjV] morae and the [CV] counterparts is more similar to the durational

[1] Although the vowel [ɯ] in [ɾja.kɯ] in Standard Japanese is typically treated as an unrounded back vowel [ɯ], Nogita, Yamane, and Bird's (2013) ultrasound study revealed that this vowel pronounced in careful speech by linguistically naïve native SJ speakers is central or rather front with lip rounding and can be with unambiguous lip protrusion, so its actual realizations are [ɯ - ʏ]. I follow their recommended symbol [ɯ] or /ɯ/.

difference between Russian [CjV] syllables and the [CV] counterparts than to the difference between Russian [C^{j}V] syllables and the [CV] counterparts. I consider that a cross-linguistic comparison would shed light on this debate, especially with a language like Russian that has both uncontroversial palatalized /C^{j}/s, as in [s^{j}est^{j}] 'sit down' as well as uncontroversial consonant clusters /Cj/s, as in [sjest^{j}] 'eat up' (examples are from Kenstowicz, 1994, p. 42).

In the second part of this paper, I examine whether [j] in [Cjʉ̥C] (C = voiceless consonant) blocks [ʉ] devoicing in SJ. This additional experiment is designed based on feedback I received at a conference. SJ has a common rule that /ʉ/ between voiceless consonants are regularly devoiced; i.e. /CʉC/ → [Cʉ̥C]. However, if [j] in [Cj] is found to block /ʉ/ devoicing, I can argue that /ʉ/ in /CjʉC/ is not between voiceless consonants and that [j] is an individual (voiced) consonant. I will examine this with nonsense words.

I am aware that there is also another argument that [j] is part of a complex nucleus /iV/ (/ / = non-syllabic/moraic), as some linguists (e.g., Hashimoto, 1984) propose. However, the goal of this study is only to reject the /C^{j}/ hypothesis, so I will not discuss the complex nucleus analysis.

1 Duration of Japanese [Cj] sequences

1.1 Brief introduction to Standard Japanese phonology

1.1.1 The Japanese onset phoneme inventory

This paper adopts Vance's (2008) and Larson-Hall's (2004) 19 Standard Japanese (SJ) onset inventory, as shown in Table 1, in which the so-called loanword sounds /ɕ, tɕ, ʦ, dʑ, ɸ/ are included.[2]

	bilabial	dental	alveolar	alveopalatal	palatal	velar	glottal
plosive	p	t				k	
	b	d				g	
affricate		ʦ		tɕ			
		dʑ		dʑ			
fricative	ɸ	s		ɕ			h
nasal	m	n					
liquid			ɾ				
glide	w				j		

Table 1 SJ 19-onset-phoneme inventory

Note: I excluded the consonants that occur only in coda positions or in the special/dependent mora positions, as they are irrelevant to my analysis.

[2] Grenon (2005) states that /dz, dʑ/, rather than /z, ʑ/, are underlying representations (UR) in SJ from a phonetic point of view. In this system, all the fricatives lack voicing contrasts while all the stops and affricates have them.

1.1.2 The Japanese palatal consonants [ɕ, tɕ, dʑ]

How the so-called SJ palatal(ized) consonant series are treated and whether the loanword stratum is included have considerably varied in the analyses of the SJ consonant inventory. Since different assumptions of the onset inventory could lead to different conclusions regarding [Cj]s, deciding on an inventory is crucial.

Based on the [Cj]=/Cj/ assumption, Vance (2008) states that 11 of the 19 consonants can precede /j/, namely, /pj, bj, ɸj, mj, tj, dj, nj, ɾj, kj, gj, hj/, whereas sibilants and glides cannot, as sequences like */sj, ʦj, ɕj, tɕj, dʑj, ʣj, wj, jj/ are not attested. Like Vance (2008) and Larson-Hall (2004), I regard the so-called palatal(ized) series [ɕ, tɕ, dʑ] as single phonemes /ɕ, tɕ, dʑ/ as opposed to the other 11 /pj, bj, ɸj, mj, tj, dj, nj, ɾj, kj, gj, hj/ for two reasons related to articulation and phonotactics.

First, Hall (2000) proposes two types of palatalization: *'true' palatalization* which adds [-back] (tongue fronting), and *nonanteriorization* which adds [-anterior, +distributed] (shift from alveolar to postalveolar). For example, (1) and (2) show how /t, d, s, z, n, l/ change based on these two types of palatalization.

(1) /t, d, s, z, n, l/ → [tʲ, dʲ, sʲ, zʲ, nʲ, lʲ] by true palatalization

(2) /t, d, s, z, n, l/ → [ʧ, ʤ, ʃ, ʒ, ɲ, ʎ] by nonanteriorization

Since SJ [ɕ, tɕ, dʑ], as in [ɕakai] 'society', [otɕa] 'green tea', and [dʑama] 'interruption' are not realized as *[sʲ, tʲ, dʲ], SJ [ɕ, tɕ, dʑ] are better treated as consonants with *nonanteriorization*.

Second, in SJ, [ɕ, tɕ, dʑ] contrast with the non-palatal counterparts [s, t, d] when preceding the vowel [i] (i.e. [ɕi, tɕi, dʑi] vs. [si, ti, di]). Other non-palatal consonants like [p, m, n, k, h], however, never contrast with [pj, mj, nj, kj, hj] before /i/, nor does *[ji] contrast with [i] (i.e. for example, [pi, ni, ki, i] vs. *[pji, nji, kji, ji]). In more details, Matsuzaki (1993) states that the /ti, di/ versus /tɕi, dʑi/ contrasts are unanimously accepted by linguists. In fact, there are (near) minimal pairs like /ti:/ ティー 'tea' versus /tɕii/ 地位 'status' as well as /diNkɯsɯ/ ディンクス 'double income no kids' versus /dʑiNkɯsɯ/ ジンクス 'jinx'. According to Nogita (2010), /si/ and /ɕi/ are marginally contrastive, as in /midʑuhaɕi paɾɯsi(:)/ 水橋パルスィ (a name of a game character). There are also /ʦi, dʑi/ as opposed to /tɕi, dʑi/, as in /eɾiʦiN/ エリツィン 'Yeltsin' (a Russian name) (Vance, 2008, p. 84) and /dʑi:dʑi:/ ズィー・ズィー (a name of a manga character).

It should also be noted that, [ɕ, tɕ, dʑ] before another front vowel [e] (i.e. [ɕe, tɕe, dʑe]) exist stably in loanwords, /tɕesɯ/ チェス 'chess', for example, whereas that of /je/ is not as stable (Takayama, 2005); /je/ is often replaced with /ie/, as in /ieɾo:/ イエロー 'yellow' rather than /jeɾo:/. Moreover, *[Cje]s, such as *[pje, kje], are not allowed (although some linguists acknowledge the existence of [hje, nje] (Matsuzaki, 1993)). All this suggests that [ɕe, tɕe, dʑe] behave differently from [je] and [Cje].

For these reasons, [ɕ, tɕ, dʑ] are phonologically treated as /ɕ, tɕ, dʑ/, but not as /sj, tj, dj (dʒj)/ in this paper. This means that the so-called *yōon* 拗音 (morae written with one regular-sized *kana* letter along with another small kana letter, such as <きゃ> (/kja/) or <しゃ> (/ɕa/)) in the regular Japanese term is not necessarily equivalent to /CjV/s. In other words, although [ɕV, tɕV, dʑV]s are *yōon* morae, I treat them as CV morae in this paper. Furthermore, In §2.3.4, I will provide further evidence that at least [dʑV]s behave like CV morae in terms of duration, as opposed to [CjV]s.

1.2 Previous studies of durations in SJ [CjV]

Durations in SJ [CjV]s have been discussed in previous studies. For example, Kida (1998) reports that phonetically [kja] is longer than [ka], which in turn is longer than [a]. Likewise, Yamaoka (2008) finds that SJ [j] in [ja] is longer than [j] in [bja, pja, mja], which in turn is longer than [j] in [kja, gja]. Note that these two authors do not provide acoustic data in their studies.

According to Parker (2012), one criterion distinguishing complex onsets and palatalized consonants is as follows: if [Cj] is a complex onset, duration of [j] is approximately as long as that of a single unambiguous onset glide. Based on this criterion, one might be inclined to the palatalized consonant hypothesis /C^{j}/ given that SJ [j] in [CjV] is shorter than [j] in [jV]. However, SJ is a mora-timed language, in which each mora is isochronous in SJ speakers' psychological timing units (but acoustically not necessarily isochronous) (Vance, 2008). In fact, according to Kohno's (1998) experiment of sentence reading, average *mora* duration in SJ is 145ms (S.D.=27.8ms, with S.D. meaning 'standard deviation'), while that of English and Spanish *syllables* are 244ms (S.D.=85.7) and 201ms (S.D.=74.0) respectively. The small S.D. values of the Japanese *morae* imply that the duration of SJ morae is relatively consistent at least in reading tasks compared to *syllables* in English and Spanish. This suggests that speakers may compress [Cj] in order to maintain the mora-timed rhythm, therefore resulting in a shorter durations of [j]. Another possible explanation can be that SJ has the [CjV]-[Ci.V] ('.' indicates a mora boundary) contrast, as in the minimal pair [kjo:] 'Buddhist sutra' versus [ki.o:] 'past illness' and [ki.jo:] 'appointment (to a position)'. Speakers may shorten [j] in [CjV] in order to make a perceptually clear contrast from [Ci.V]. In either case, in my review of the literature, I did not find studies that focused on acoustically to what extent [CjV]s are longer than [CV]s in SJ.

1.3 Experiment 1: durations of SJ [CjV]s and the [CV] counterparts

My first question is whether SJ [CjV] morae duration is longer than that of the [CV] counterparts. To examine it, I instructed native SJ speaking participants to

pronounce these morae in nonsense words embedded in a sentence, and measured the duration[3].

1.3.1 Participants

34 (11 male and 23 female) native SJ speakers, aging from 15 to 59, participated in the experiment. All the participants were reportedly born and raised in or near Tokyo: more specifically, Tokyo, Kanagawa, Saitama, Chiba, Yamanashi, and Gunma. As my experiment was conducted in Victoria, British Columbia, Canada, the participants were recruited in Victoria.

1.3.2 Stimuli and procedure

The target morae were [gʲja, nja, mja, bja, ɾʲja] versus [ga, na, ma, ba, ɾa] respectively, and the control pair was [d͡ʑa] (palatal /CV/) versus [d͡za] (non-palatal /CV/). The durations of these morae were compared. Voiced consonants were chosen for the stimuli with the purpose of avoiding a possibility of vowel devoicing/deletion when adjacent to voiceless consonants, which might affect the results. To avoid homorganicity, these morae were embedded in a 4-mora nonsense compound word [_tá-bako] when the target consonant was non-coronal, for example, [m̲atá-bak̲o], and [_ká-bako] when the target consonant was coronal, for example, [n̲ak̲á-bako]. The participants were instructed to place the phonemic pitch accent (or a high tone) on the second mora in these four-mora nonsense words; [bako] is from a real word 'box', and in real compound words, the mora before [bako] is the default accented mora (e.g. [omot͡ɕá-bako] 'toy box'), so the place of the accented mora in [_tá-bako]/[_ká-bako] is natural to SJ speakers. These nonsense words were embedded in the sentence [watad͡ʑitat͡ɕiwa _ desɯ] 'We are _.'

These stimulus sentences were printed on a sheet with regular Japanese orthography (see Appendix for the stimulus sentences). The participants read aloud each sentence three times in natural speed and only the second token was analyzed. They were encouraged to practice all the nonsense words once beforehand to familiarize themselves with the words, which also allowed me to correct their obvious mistakes if any. Recording was done in the soundproof booth in the UVic Phonetics Lab with a microphone, SONY ECM-MS908C, and the Software Audacity set at 44100Hz and 32-bit float.

[3] In my analysis, I use the term 'mora' rather than 'syllable' since the necessity of the syllable nodes in the prosodic hierarchy in SJ is debatable (e.g., Labrune, 2012). The difference between morae and syllables is that a mora counts a coda consonant and the second half of a long vowel/diphthong; as an example, the /CCV:C/ syllable has three morae, /CCV/ + /:/ + /C/.

1.3.3 Data analysis

Since it is difficult to spot the boundary between [j] and the following vowel, the durations of the whole morae were measured. In addition to the absolute duration, the ratio between the target mora and the following mora (e.g., the ration between [gja] and [tá] in [gjatábako]) was compared in consideration of the speech rate. For morae with stop/affricate/flap onsets, the measurement points were from the beginning (zero crossing) of the closure to the end of the last pitch pulse. For nasal onsets, the beginning points were the beginning of the first pitch pulse that shows a sudden change of the waveform pattern (see Figure 1).

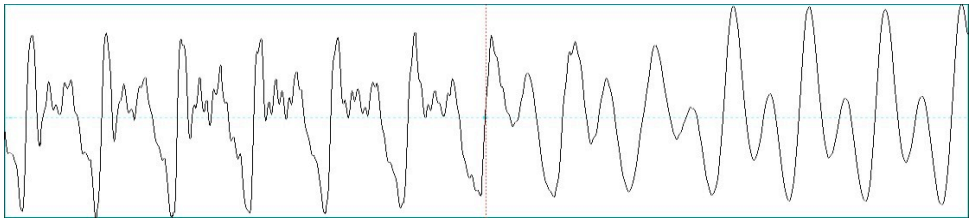

Figure 1. The measurement point of the beginning of a nasal consonant; the vertical red dot line indicates the measurement point.

The analysis was done using the phonetic software PRAAT. All the data were measured twice, and the consistency across measurement 1 and 2 was 86.5%. Despite the relatively low consistency, [Cja]s were longer than [Ca]s by 12ms to 25ms on average, as shown in Table 2 in §2.3.4, so even if measurement points in some tokens were wrong by one pitch pulse (roughly 3ms to 7ms depending on the gender), the overall results would still be the same.

1.3.4 Results and discussion

Table 2 shows the average durations and average ratios.

Mora	dza	dʑa	ga	gja	na	nja	ɾa	ɾja
Duration[a]	147	144	130	142	134	156	115	132
p-value[b]	p=0.28		*p=0.00062		*p=3E-08		*p=6E-05	
Duration of tá/ká	138	141	147	140	143	144	147	143
p-value	p=0.39		*p=0.042		p=0.60		p=0.15	
Ratio %	109%	104%	89%	103%	95%	109%	79%	93%
p-value	p=0.18		*p=4E-30		*p=1E-07		*p=2E-06	

Mora	ma	mja	ba	bja
Duration[a]	135	160	136	161
p-value[b]	*p=1E-08		*p=6E-08	
Duration of tá/ká	142	138	140	138
p-value	p=0.15		p=0.29	

Table 2 Average durations of unaccented [C(j)a] and the following [tá/ká], and average ratios between [C(j)a] and the following [tá]/[ká] in Standard Japanese
Note:[a] The durations are given in milliseconds. [b] p-values were calculated by the two-tailed paired t-test between [C(j)a]s and their [Ca] counterparts. Asterisk* indicates a significant difference.

The [gja, nja, mja, bja, ɾja] morae were consistently longer than corresponding [ga, na, ma, ba, ɾa]; note that according to the [C(j)a]-[tá/ká] ratios, in only 20 (out of 170) cases, [CV] was longer than its [CjV] counterpart of the same speaker (6 in [ga-gja], 5 in [na-nja], 2 in [ma-mja], 3 in [ba-bja], 4 in [ɾa-ɾja]). Overall, both absolute durations and the [C(j)a]-[tá/ká] ratios of [gja, nja, mja, bja, ɾja] were highly significantly longer than the [ga, na, ma, ba, ɾa] counterparts. In contrast, between [ʣa] and [ʥa], unexpectedly, the non-palatal [ʣa] was slightly longer than the palatal [ʥa], but the difference was not significant. Since the non-palatal versus palatal pair [ʣa]-[ʥa] did not show a difference in duration, at least in SJ, the generalization that palatal consonants are longer than the non-palatal counterparts was not observed. In other words, I can argue that the fact that [gja, nja, mja, bja, ɾja] are significantly longer than [ga, na, ma, ba, ɾa] is not due to the nature of palatal consonants, but because [gja, nja, mja, bja, ɾja] consist of three segments /CCV/ while [ga, na, ma, ba, ɾa] as well as [ʣa, ʥa] consist of two segments /CV/.

1.3.5 Other findings that support previous studies

One result of this experiment is that [gja] was longer than [ga] only by 12ms, while [mja, bja] were longer than [ma, ba] by 25ms. This is consistent with Yamaoka's (2008) statement that [j] in [pj, bj, mj] is longer than [j] in [kj, gj]. The occasions that [mja, bja] failed to be longer than [ma, ba] were two and three times respectively, while the occasions that [gja] failed to be longer than [ga] were as many as six times. In other words, [mja, bja] showed long durations more consistently than [gja] did.

I suspected that there may be compensatory shortening in the following mora [tá/ká] to maintain the mora-timed rhythm, that is, if a [CjV] mora is long, speakers might shorten the following mora to keep the same duration of the whole word. According to Table 2, except for [ká] after [nja, na], the mora [tá/ká] after [Cja] was slightly shorter than [tá/ká] after [Ca]. However, a statistically significant shortening was found only in [tá] after [ga, gja]. Moreover, the difference between [tá] after [gja] and [tá] after [ga] was only 7ms on average, which may not be perceptually salient. Therefore, there was no solid evidence for compensatory shortening. Based on my perceptual impression, the nonsense words containing [Cja] sometimes sounded slower than those without [Cja]. Given this impression, I also suspected a global slowdown by lengthening the following mora in order not to make the [CjV] mora sound deviant from the surrounding morae in duration as an alternative strategy for the mora-timed rhythm. However, as Table 2 shows, this was not the case either. That is, neither compensatory shortening of the following mora nor a global slowdown was observed in this experiment. These results agree with Warner and Arai (2001), who do not find any compensation related to the mora-timed rhythm either. This means that my perceptual impression that the nonsense words containing [Cja] sounded slower than those without [Cja] is only because of the longer duration of [Cja]. Still, Warner and Arai (2001, p. 1149) find a high correlation between duration and the number of morae in spontaneous speech by 11 native Japanese (not necessarily SJ) speakers (the r-value ranging from 0.701 to 0.931). Thus, there is a possibility that short duration in [j] in [CjV] comes from an attempt to keep the mora-timed rhythm.

1.4 Intermediate conclusion

To the question of whether SJ [CjV] morae are phonetically longer than their [CV] counterparts, the answer is yes. [CjV]s are highly significantly longer than the [CV] counterparts. Since the palatal [dʑ] was not longer than the non-palatal counterpart [dz], the reason that [CjV]s are longer than corresponding [CV] is not because of the long duration of palatal consonants. Instead, this suggests a possibility that [CjV]s consist of three segments.

2 Duration of Russian [C^j] and [$C^{(ɣ)}$]

2.1 The second research question

In the last section, I concluded that palatal consonants are not phonetically longer than their non-palatal counterparts. However, only the coronal affricate [dʑ] and [dz] were examined. In this section, I examine whether consonants with true palatalization (see §2.1), such as [n^j, m^j, b^j, r^j], do not show longer durations than corresponding [n, m, b, r]. If [n^j, m^j, b^j, r^j] are clearly longer than [n, m, b, r] respectively, then it can be concluded that SJ [nj, mj, bj, ɾj] are longer than [n, m, b, ɾ] because of the longer duration of palatalized consonants. In other words, SJ [nj, mj, bj, ɾj] can be interpreted as single palatalized consonants /n^j, m^j, b^j, $ɾ^j$/. In contrast, if [n^j, m^j, b^j, r^j] are not longer than [n, m, b, r], it can be another piece of evidence that SJ [nj, mj, bj, ɾj] are consonant clusters /nj, mj, bj, ɾj/. In order to examine the duration of [C^j]s, I investigate Russian, which has unambiguous contrastive palatalized consonants /C^j/s (also called soft consonants, мягкие согласные) as opposed to plain/velarized consonants /$C^{(ɣ)}$/s (also called hard consonants, твердые согласные). In this paper, I omit the velarized symbol /^{ɣ}/ for hard consonants.

2.2 Brief introduction to Russian phonology

Russian has phonologically contrastive palatalized consonants and plain/velarized counterparts, as in [m^{j}at^{j}] 'to rumble', [mat] 'checkmate', [m^{j}at] 'rumpled', and [mat^{j}] 'mother' (Kenstowicz, 1994, p. 41). Table 3 shows the Russian consonant inventory from Padgett (2003b).

	labial		dental		post-alveolar		palatal		velar	
plosive	p	p^{j}	t	t^{j}					k	k^{j}
	b	b^{j}	d	d^{j}					g	g^{j}
fricative	f	f^{j}	s	s^{j}	ʃ	ʃ^{j}:			x	x^{j}
	v	v^{j}	z	z^{j}	ʒ					
affricate			ʦ			ʧ^{j}				
nasal	m	m^{j}	n	n^{j}						
liquid			l	l^{j}						
			r	r^{j}						
glide							j			

Table 3 Russian consonant inventory from Padgett (2003b)

Padgett (2003a) mentions that if a language has palatalized consonants, plain counterparts tend to be velarized for the sake of perceptually clear distinctiveness. For example, Kochetov (2002) observes that Russian [p^{j}] in [ap^{j}á] involves tongue body fronting and raising, while [p] in [apá] involves partial velarization (or pharyngealization), that is tongue body backing but without tongue body raising. Still, according to Litvin's (2014) ultrasound study, clear presence of velarization/pharyngealization in plain consonants is debatable. In any case, palatalized-plain/velarized contrasts are [-/+back] contrasts (Padgett, 2003a). I will examine durations of [-/+back] consonants.

In addition, Russian [C^{j}]s also phonologically contrast with consonant cluster [Cj]s as in [s^{j}est^{j}] 'sit down' versus [sjest^{j}] 'eat up' (Kenstowicz, 1994, p. 42). This contrast is what Japanese lacks. Besides, Russian also has a syllabicity contrast, such as [di.a] versus [dja] (Padgett, 2008), just as Japanese does.

2.3 Experiment 2: durations of Russian [C]s, [C^{j}]s, and [Cj]s in unstressed positions

2.3.1 Participants

I recruited eight (3 male and 5 female) native Russian speakers, aging from 18 to 29, and reportedly born and raised in Russia (European Russia, the South-West area, and Perm), Kazakhstan, and Belarus. Again, all the participants were recruited in Victoria, British Columbia, Canada.

2.3.2 Stimuli, procedure, and analysis

The three groups of target syllables were [na, ma, ba, ra], [n^{j}a, m^{j}a, b^{j}a, r^{j}a], and [nja, mja, bja, rja], which were the counterparts of the morae used in the Japanese experiment respectively. In the Russian experiment, the velar [g^{j}a] was discarded in analysis (although they were also recorded). This is because in Russian, while the hard-soft contrast in velars before /a, o, u/ is present in a fair number of loanwords, there is some dispute whether velars are contrastive in palatalization (Padgett, 2003b) and the speakers may not be familiar with it as well as the other palatalized consonants.

Like in the Japanese experiment, the stimulus syllables were embedded in the nonsense word template /_kábako/ or /_tábako/. Again, to avoid homorganicity, the template was /_tábako/ when the target consonant was non- coronal (e.g., /matábako/) and /_kábako/ when the target consonant was coronal (e.g., /nakábako/)[4]. The participants were instructed to place the phonemic stress on the second syllable in order to replicate the Japanese experiment[5]. These nonsense words were embedded in the sentence /mi _/ (or /mɨ _/, depending on the phonological interpretation) 'We are _.'

The stimuli were printed on a sheet with regular Cyrillic Russian orthography (see Appendix for the stimuli). The recording procedure and data analysis were done in the same way as in the Japanese experiment. The only difference is that in the Russian experiment, participants read aloud each sentence five times in natural speed and the middle three tokens were analyzed. This is for compensating my limited access to Russian speakers in Victoria.

2.3.3 Results and discussion

Table 4 shows the average durations and average ratios regarding [CV]s and [C^{j}V]s in unstressed syllables. Table 5 shows those of [CV]s and [CjV]s.

[4] In the section of the Japanese experiment, I used the phonetic brackets [ ], as whether [Cj] is phonologically /C^{j}/ or /Cj/ had not been decided yet. In the Russian experiment, however, I use the phonemic brackets / /, since the status of the Russian consonants at issue is already uncontroversial. In addition, unstressed /a/ and /o/ are neutralized so that the orthographically indicated vowels in the stimuli can be different from the actual realizations. Therefore, using / / for the Russian stimulus sentences is more suitable.

[5] When pitch is involved, I use the term *accented*, and otherwise *stressed*. Since the Japanese pitch accent involves pitch, I call it *accented*. Since the Japanese pitch accent system can be categorized as a type of the tone system rather than the stress system (Hyman, 2006), the Russian experiment cannot exactly be a replication of the Japanese one in terms of stress/accent patterns.

Syllable	na	n^ja	ma	m^ja	ba	b^ja	ra	r^ja
Duration[a]	130	139	142	147	150	157	109	106
p-value[b]	*p=0.012		p=0.147		p=0.228		p=0.450	
Duration of tá/ká	197	197	198	193	192	189	196	194
p-value	p=0.943		p=0.420		p=0.620		p=0.608	
Ratio %	68%	74%	75%	79%	82%	86%	58%	56%
p-value	*p=0.006		p=0.066		p=0.365		p=0.480	

Table 4 Average durations of [C(^{j})a] and the following [tá/ká], and average ratios between [C(^{j})a] and the following [tá]/[ká] in Russian

Note:[a] The durations are given in milliseconds. [b] p-values were calculated by the two-tailed paired t-test between [Ca]s and their [C(^{j})a] counterparts. * indicates a significant difference at the $p<0.05$ level.

Syllable	na	nja	ma	mja	ba	bja	ra	rja
Duration	130	196	142	200	150	197	109	177
p-value	*p=2E-08		*p=3E-10		*p=5E-06		*p=2E-11	
Duration of tá/ká	197	207	198	203	192	198	196	208
p-value	p=0.058		p=0.31		p=0.43		p=0.10	
Ratio %	68%	97%	75%	103%	82%	103%	58%	87%
p-value	*p=6E-10		*p=5E-07		*p=4E-06		*p=2E-09	

Table 5 Average durations of [C(j)a] and the following [tá/ká], and average ratios between [C(j)a] and the following [tá]/[ká] in Russian

According to the [C^j]-[C] comparison in Table 4, unlike the case of highly significant differences in all the Japanese [Cj]-[C] pairs, these Russian [C^j]-[C] pairs did not show significant differences in both absolute durations and the [C(^{j})a]-[tá/ká] ratios, except that [n^j] was significantly different from [n] in both the absolute duration ($p<0.05$) and in the ratio ($p<0.01$). It should be noted, however, that [n^ja] is longer than [na] only by 9 ms in the absolute duration and 6% in ratio on average, while the Japanese [nja, mja, bja, ɾja] were longer than corresponding [na, ma, ba, ɾa] by 17-25ms in duration and 14-21% in ratio on average[6]. Likewise, Russian [m^j] was almost significantly longer than [ma] in the [m^ja]-[tá] ratio ($p=0.066$) (not significant in the absolute duration), but only by 4%. The differences in these two cases may not be perceptually salient. Therefore, there is no clear evidence indicating that palatalized consonants are longer than plain counterparts. Finally, I want to point out that while Japanese [CjV]s are impressionistically longer or slower than [CV]s, this is not the case with Russian [C^jV]s.

[6] Regarding the difference between Russian [n^ja] and [na] as opposed to the difference between Japanese [nja] and [na], one reviewer asked whether the difference in Japanese was statistically bigger than those in Russian. The two-tailed two-sample t-test showed that the difference in Japanese was significantly bigger than that in Russian ($p=0.013$ in duration and $p=0.005$ in ratio).

The question which deserves further discussion is why only [n^{j}a] was significantly longer than [na]. This may be because the preceding segment is the particular vowel /i/. Since the place of articulation of /i/ and that of /n^{j}/ is close, there may be partial assimilation in the end part of /i/, which ends up with slight lengthening of /n^{j}/. As a piece of evidence, while Figure 2 shows that the boundary between /i/ and /n/ is very clear, Figure 3 shows that the boundary between /i/ and /n^{j}/ by the same speaker is not as clear, and all the speakers show the same tendency. These data seem to show this assimilation in the /in^{j}/ sequence.

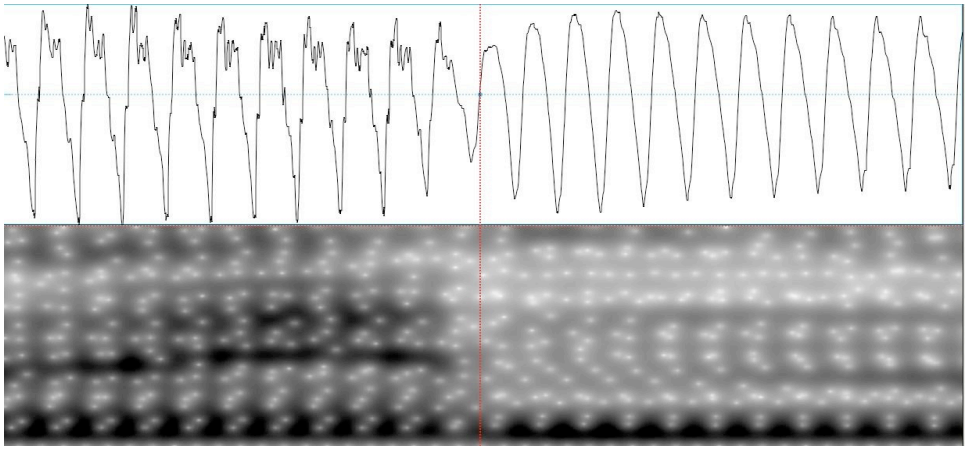

Figure 2 Boundary between /i/ and /n/ in /mi nakábako/

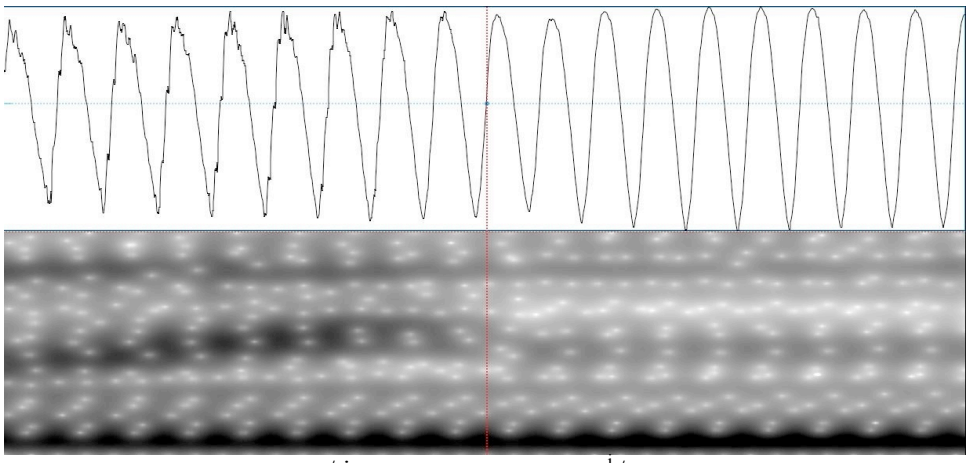

Figure 3. Boundary between /i/ and /n^{j}/ in /mi n^{j}akábako/

As for the [CjV]-[CV] comparison shown in Table 5, it is more straightforward. Table 5 indicates that [CjV]s are highly significantly longer than the [CV] counterparts. In the [C(j)a]-[tá/ká] ratios, only in one occasion ([mja- ma]) (out of the total 96) that [CjV] failed to be longer than the [CV] counterpart for the same speaker. In the absolute durations, only two occasions (both in [bja-

ba]) showed that [CjV] failed to be longer. The differences in the absolute durations between [CjV]s and [CV]s are 47-68ms and those of the [C(j)a]-[tá/ká] ratios are 21%-29%. Compared to the standard Japanese (SJ) cases (17-25 ms and 14-21%), Russian [CjV]s are longer than the [CV] counterparts more consistently and by larger degrees. This can be analyzed that the phonological contrasts between /CjV/s and the /C^{j}V/ counterparts need to be perceptually salient in Russian in order to maintain the contrasts. In contrast, SJ lacks the phonological /CjV/-/C^{j}V/ contrast, and therefore there is no motivation to lengthen [CjV]s. Moreover, SJ [CjV]s have to contrast with [Ci.V]s and have to keep the mora-timed rhythm. Thus, [CjV] with a long duration is rather not preferable. Therefore, the above analysis suggests that the fact that Japanese [CjV]s are longer than [CV]s less consistently and by smaller degrees than Russian counterparts cannot be the evidence that Japanese [Cj]s are phonologically /C^{j}/s.

As for the following [tá/ká] syllables, there were no significant differences between [tá/ká] following [CV]s and [tá/ká] following [C^{j}V]s, or between [tá/ká] following [CV]s and [tá/ká] following [CjV]s, suggesting that neither global slow down nor compensatory shortening is observed. This is the same as Japanese.

2.4 Experiment 3: durations of Russian [C]s, [C^{j}]s, and [Cj]s in stressed positions

2.4.1 Possible stress effect

There is a possibility that the syllables in the experiment above were pronounced as unstressed syllables and their length differences between /C^{j}V/s and /CV/s may have been therefore reduced. To test this possibility, in this section, I examine Russian syllables in a stressed position.

2.4.2 Participants, stimuli, and procedure

The participants in this experiment were the same as experiment 2. The target syllables were also [na, ma, ba, ra] versus [n^{j}a, m^{j}a, b^{j}a, r^{j}a] versus [nja, mja, bja, rja], except that they were stressed. In experiment 2, the stimuli used were replication of the Japanese nonsense words, which should have sounded very foreign for the Russian participants. In this experiment, the stimulus nonsense words were designed to be more Russian-like. The stressed target syllables were embedded in /__tap/ or /__kap/, for example, /n^{j}ákap/ and /m^{j}átap/. Again, /t/ and /k/ were alternated to avoid homorganicity. These stimulus words were embedded in /ona hot^{j}ela __ / 'She wanted __.' This time, the vowel immediately before the target syllable is /a/ instead of /i/ in order to avoid potential assimilation with the following palatalized consonant. The participants pronounced each sentence five times and the middle three were analyzed. Only the /ta/ or /ka/ was measured in the second syllable, and the coda /p/ was excluded.

2.4.3 Results and discussion

Table 6 shows the average durations and average ratios regarding [CV]s and [C^{j}V]s in stressed syllables. Table 7 shows those of [CV]s and [CjV]s.

Syllable	na	n^{j}a	ma	m^{j}a	ba	b^{j}a	ra	r^{j}a
Duration[a]	188	195	196	211	208	215	141	153
p-value[b]	p=0.055		*p=0.0002		*p=0.025		*p=0.005	
Duration of tá/ká	143	147	145	149	148	148	149	151
p-value	p=0.177		p=0.163		p=0.848		p=0.667	
Ratio %	136%	138%	142%	150%	148%	155%	98%	104%
p-value	p=0.533		p=0.061		*p=0.022		p=0.106	

Table 6 Average durations of stressed [C(j)á] and the following [ta/ka], and average ratios between [C(j)á] and the following [ta]/[ka] in Russian

Note.[a] The durations are given in milliseconds. [b] p-values were calculated by the two-tailed paired t-test between [Ca]s and their [C(j)a] counterparts. * indicates a significant difference at the p<0.05 level.

Syllable	na	n^{j}a	ma	m^{j}a	ba	b^{j}a	ra	r^{j}a
Duration[a]	188	250	196	258	208	255	141	213
p-value[b]	*p=4E-11		*p=1E-10		*p=1E-08		*p=7E-10	
Duration of tá/ká	143	157	145	147	148	148	149	153
p-value	*p=0.007		p=0.554		p=0.859		p=0.289	
Ratio %	136%	166%	142%	185%	148%	182%	98%	145%
p-value	*p=3E-05		*p=6E-08		*p=1E-05		*p=2E-08	

Table 7 Average durations of stressed [C(j)á] and the following [ta/ka], and average ratios between [C(j)á] and the following [ta]/[ka] in Russian

As shown in Table 6, interestingly, in stressed syllables, [C^{j}V]s are significantly longer than [CV]s in absolute durations, except that [n^{j}á] is only marginally significantly longer than [ná]. Moreover, the difference is 7ms to 15ms on average, so at least the [m^{j}á]-[má] and [r^{j}á]-[rá] differences might be perceptually noticeable (and according to my perceptual impression, [m^{j}á] sometimes sounded slightly longer than [má]). However, as mentioned above, absolute durations vary depending on the speech rate, so that ratios with the adjacent syllable would be more reliable. In fact, the [C^{j}á]-[ta/ka] ratios and the [Cá]-[ta/ka] ratios were not significantly different, except that [b^{j}á] was significantly longer than [bá] at the p<0.05 level. Also, the [m^{j}á]-[má] difference was marginally significant. However, the statistically significant difference between [b^{j}á] and [bá] was only 7%. Recall that the Japanese [nja, mja, bja, ɾja] were longer than [na, ma, ba, ɾa]

by 14-21% on average even in an unaccented low tone position with the mora-timed rhythm restriction[7]. This means that even the significant values in stressed syllables in Russian were much less obvious than the Japanese values. What is intriguing is that this time, the stressed [n^{j}á]-[ná] difference was not significant. This suggests that the significance in the unstressed [n^{j}a]-[na] difference in the previous experiment was likely because of partial assimilation of the preceding /i/ à [in^{j}]. In other words, the significance was only accidental. So I conclude that in Russian, only palatalized bilabial consonants [m^{j}] and [b^{j}] in stressed syllables tend to be slightly longer than the plain/velarized counterparts [m] and [b], but the difference is much less than the difference between Japanese [mj, bj] and [m, b].

Table 7 shows, unsurprisingly, that [CjV]s were significantly longer than the [CV] counterparts. In the [C(j)á]-[ta/ka] ratios, in five cases (1 in [mja], 2 in [bja], and 2 in [nja]) out of 96, [CV] was longer than the [CjV] counterpart by the same speaker, and in the absolute durations, only in one case (in [bja]) [CV] was longer. The differences in the absolute durations between [CjV]s and [CV]s were 47-72 ms and those of the [C(j)a]-[tá/ká] ratios were 30%-47%. These data indicate that in stressed positions the differences between [CjV]s and [CV]s become larger than in unstressed positions. Still, it is interesting that in Russian [CjV]s are occasionally shorter than [CV]s, suggesting that the fact that Japanese [CjV]s occasionally become shorter than [CV]s does not mean that Japanese [Cj]s are phonologically /C^{j}/s. Again, compensatory shortening of the following syllable or global slowdown is not observed, except that stimuli with [njá]s show global slow down.

2.5 Intermediate conclusion

My experiments indicate that while in both Japanese and Russian [CjV]s were highly significantly longer than the [CV] counterparts, Russian [C^{j}V] versus [CV]s did not show the same pattern. This suggests that Japanese [CjV]s are more similar to Russian /CjV/s than to Russian /C^{j}V/s. My experiments also show that Russian stressed bilabial [C^{j}]s may be slightly longer than their [C] counterparts, but the difference is not as significant as the difference between Japanese unaccented [Cj]s and [C]s. These phenomena seem to be better explained by suggesting that Japanese [CjV]s consist of three segments, rather than that longer durations are the nature of palatalized consonants. At least, there is no duration-related evidence that Japanese [Cj]s are closer to /C^{j}/s than they are to /Cj/s.

[7] Again, to answer one reviewer's question whether the difference between Japanese [bja] and [ba] (25ms, 20%) was statistically more significant than that of Russian [b^{j}a] and [ba] (7ms, 7%), I ran the t-test, which showed that the difference in Japanese was significantly bigger ($p<0.001$ for duration and $p=0.002$ for ratio).

3 Possible blocking of high vowel devoicing in Standard Japanese

3.1 Background and research question

When presenting the research on the Japanese [Cj] at the annual conference of The Phonological Society of Japan in 2013, I received feedback that if [j] blocks vowel devoicing, it can be evidence that [Cj] is a consonant cluster. Based on this feedback, I designed an additional experiment.

My research question is whether [j] in [Cj] blocks High Vowel Devoicing/Deletion (HVD) in SJ. As a common generalization, the short high vowels /ʉ/ and /i/ between voiceless consonants are typically devoiced/deleted in SJ, i.e. [ʉ] and [i] in /C̥ʉC̥, C̥iC̥/ (C=voiceless consonant) become [C̥ʉC̥, C̥iC̥] or [CC, CC]. If one of the consonants adjacent to [ʉ] or [i] is voiced, as a common generalization, HVD does not occur. Therefore, in [jʉC̥] contexts, /ʉ/ is not likely to be devoiced/deleted since [j] is a voiced consonant. Likewise, if [j] in [CjV] is the same consonant as [j] in [jV], the [CjʉC̥] contexts would block HVD of /ʉ/. Recall that since *[ji] and *[Cji] are disallowed in SJ (see §2.1), *[CjiC̥] does not occur, so I examine only [CjʉC̥].

In my review of the literature, only Kondo (2000) attempted to examine whether [Cj] is a complex onset or a palatalized consonant by observing HVD blocking in [CjʉC̥] contexts in real words, such as /okjʉpe:ɕoN/ ‘occupation’ (p. 135). Her finding shows that [j] in [CjʉC̥] blocked HVD 46 times out of 56 occasions in the word reading task by four native SJ speakers. If her results can be generalized, [Cj] would be an unambiguous complex onset because it can be explained that the voiced [j] changes the environment where [ʉ] is sandwiched between voiceless consonants. However, Kondo also mentioned that she used only real words, but [CjʉC̥] does not frequently occur in the actual vocabulary. Moreover, while an unaccented /ʉ/ in /C̥ʉS/ (S=voiceless stop) causes HVD nearly 100% of the time in SJ (Fujimoto & Kiritani, 2003), when /ʉ/ bears a phonemic accent nucleus (/C̥ʉC̥/) or when there are two consecutive devoiceable morae (/C̥ʉC̥ʉC̥/ etc.), it is not always devoiced (Kondo, 2000). In Kondo’s (2000) experiment, the words had environments where HVD does not always occur, such as /kjʉpʉra/ ‘cupra’ with an accented /ʉ/ (p. 135). In fact, Kondo (2012, in personal communication) said that her observation was preliminary and cannot be generalized. Furthermore, contrary to Kondo (2000), according to Shinohara (2012), in prescriptive Japanese, /Cjʉ(C)/ morae (/kjʉ, hjʉ, pjʉ/) are supposed to be devoiced in order to make them sound fluent. Since Kondo’s (2000) finding is not yet conclusive, whether [j] in [CjʉC̥] blocks HVD or not by linguistically naïve native SJ speakers is still unknown. To fill in this gap, I conducted experiment 4.

3.2 Experiment 4: HVD in [CjʉC] contexts

3.2.1 Participants, stimuli and procedure

The participants were the same group that participated in Experiment 1. The target morae were the [Cjʉ] forming [kjʉ, pjʉ, tjʉ, ɸjʉ, hjʉ], the [Cʉ] counterparts forming [kʉ, pʉ, tʉ, ɸʉ] (since *[hʉ] is phonotactically prohibited, [hjʉ] lacks its counterpart). Note that [tjʉ, ɸjʉ] and [tʉ] occur only in recent loanwords, while [kʉ, pʉ] occur in native words, so a word stratum difference may affect the results (see Table 8 below). I also included other [Cʉ] morae written with two kana letters, specifically [ɕʉ, tɕʉ] (シュ, チュ), in order to examine a possibility that blocking of HVD is not due to pure phonology, but due to orthographic influence. I also compared both [Ci] morae that occur in native Japanese words, specifically [hi, tɕi], and [Ci] morae that occur only in recent loanwords, specifically [ti, ɸi], in order to examine a word stratum (or orthography) influence. These target morae are embedded in four-morae nonsense compound words. Again, the first element is [_tá] when the target consonant is non-coronal and [_ká] when the target consonant is coronal to avoid homorganicity. The second element is of a real morpheme, [boɕi] 'star', [baɕi] 'bridge', or [batɕi][8] 'bee' starting with a voiced consonant to avoid a slight possibility of devoicing in the preceding vowel. The [Cjʉ] targets and their [Cʉ] counterparts are followed by the same second element to be consistent in the environment: for example, [kjʉtá-boɕi] versus [kʉtá-boɕi], and [ɸjʉtá-batɕi] versus [ɸʉtá-batɕi]. There were also six other four-mora nonsense words as distracters. The participants were instructed to put the phonemic pitch accent nucleus on the second mora. Since the default place of the accent nucleus in this type of compound words is the last syllable of the first element, the pitch accent pattern of these stimuli should be natural to native SJ speakers. These nonsense words were embedded in the same sentence as the one used in Experiment 1, *watashitachiwa _ desu* 'We are _'. *Watashitachi,* where the underlined /i/ is highly likely devoiced/deleted, could make the participants feel natural to devoice the target morae. The participants were asked to read each sentence in natural speed three times.

This experiment and Experiment 1 in §2.3.1 were done together, and the sentences in these two experiments were shuffled in random order on the same sentence list, so that the sentences in one experiment were distracters for the other experiment (see Appendix for the actual stimuli).

[8] [boɕi], [baɕi], and [batɕi] are underlyingly /hoɕi/, /haɕi/, and /hatɕi/ respectively. In compound words, the first consonants become voiced due to the *rendaku*, or sequential voicing phenomenon.

3.2.2 Data analysis

The number of occurrence of HVD in the target was counted. All the three repetitions in each stimulus sentence were analyzed, so that each target mora was pronounced 102 times (3 repetitions × 34 participants).

3.2.3 Results and discussion

Overall, 7 out of 34 participants devoiced all the target vowels in all the three repetitions, and the others did not devoice them at least once. Table 8 shows the results.

Target [C(j)ʉ]	[kʉ]	[kjʉ]	[pʉ]	[pjʉ]	[ɸʉ]	[ɸjʉ]	[tʉ]	[tjʉ]	[hjʉ]
Word stratum[a]	N/S	N/S	N/S	N/S	N/S	Fo	Fo	Fo	N/S
% of HVD[b]	91	75	84	65	98	66	65	63	67
p-value[c]	*p=0.011		*p=0.002		*p<0.001		p=0.571		

Target [Ci]	[ti]	[ɸi]	[hi]	[ʨi]
Word stratum	Fo	Fo	N/S	N/S
% of HVD	73	75	94	95

Target [Cʉ]	[ʦʉ]	[ʨʉ]
Word stratum	N/S	N/S
% of HVD	61	87

Table 8. The number of high vowel devoicing

Note .[a] N/S (native/Sino)=occurring in native Japanese and Sino-Japanese words and in recent loanwords as well; Fo (foreign)=occurring in recent foreign loanwords but not in Sino-Japanese and native vocabulary. Orange shaded boxes are morae only in recent foreign loanwords. [b] % of HVD is the percentage of vowel devoicing out of 102 tokens. [c] p-values were calculated by the two-tailed paired t-test. Asterisk* indicates a significant difference between [Cʉ] and [Cjʉ] at the $p<0.05$ level.

Between [Cjʉ]s and their [Cʉ] counterparts, [Cjʉ]s were consistently less frequently devoiced, although [tjʉ] and [tʉ], both of which are loanword morae, were not significantly different. This suggests that [j] blocks HVD to some extent. However, if I look at other morae, the loanword CV morae [ti, ɸi] were devoiced less frequently than native/Sino morae [hi, ʨi].

Frequency of HVD may be related to frequency of occurrence in the Japanese vocabulary. According to Takayama (2003), occurrences of [Cj]s are relatively few in native Japanese words compared to Sino-Japanese words, recent loanwords, and onomatopoeias. In addition, Hizume (2003) points out asymmetrical distribution of long and short vowels in Sino-Japanese words,

where [Cj]s generally precede a phonemically contrastive long /ʉː/ rather than short /ʉ/. This means that [kjʉ, pjʉ, hjʉ] (with a short vowel) are infrequent. Also, according to Otake, Hatano, Cutler, and Mehler (1993), over 70% of morae in corpora of Japanese speech are [CV]s (and the rest are [CjV]s, [C]s, [V]s, and the second half of a long vowel). In consideration of these studies, my results could also be interpreted that infrequently occurring morae [kjʉ, pjʉ, tjʉ, ɸjʉ, hjʉ, tʉ, ti, ɸi] may be less likely to be devoiced regardless of the mora structures.

Another possible interpretation is that *Yōon* [CjV]s and [CV]s, as well as recent loanword [CV]s, are written with two letters of the Japanese *Kana* syllabary, while native [CV]s are written with one letter as shown from (3) to (6):

(3) Native [CV]s [kʉ, ɸʉ, pʉ, hi, tɕi] are ク, フ, プ, ヒ, チ, respectively.

(4) [CjV]s [kjʉ, pjʉ, tjʉ, ɸjʉ, hjʉ] are キュ, ピュ, テュ, フュ, ヒュ.

(5) Loanword [CV]s [tʉ, ti, ɸi] are トゥ, ティ, フィ.

(6) *Yōon* [CV]s [ɕʉ, tɕʉ] are シュ, チュ.

Orthography might have affected the participants' productions. However, there is a case that orthography cannot explain; two-letter [tɕʉ] チュ is slightly more frequently devoiced than one-letter [pʉ] プ (87% vs. 84%). In other words, orthography does not explain the whole story.

As for [ɕʉ] and [tɕʉ], these [CV] morae were arguablly developed from Sino-Japanese words and are written with two *kana* letters (シュ, チュ) respectively. Even though both are [CV] morae with the same phonological status in terms of word stratum, [ɕʉ] was significantly ($p < 0.001$) less frequently devoiced. Interestingly HVD in [ɕʉ] was even less frequent than that in the loanword [Cjʉ] morae [ɸjʉ] and [tjʉ] although the differences were slight. All this indicates that there are many factors involved in HVD, which I will not discuss in this paper.

To sum up, /j/ in /CjʉC/ could be one of the factors of HVD blocking, but since there are other possible factors, further research is still needed.

Meanwhile, although more than 60% of the occurrences of /ʉ/ in /CjʉC/ were devoiced, it does not reject the complex onset hypothesis. For example, English /l/ and /ɹ/ after voiceless consonants are allophonically devoiced as in 'play' [plej] and 'tray' [tɹej] (McMahon, 2002, p. 65). Likewise, allophonic devoicing of /j/ in /CjʉC/ can also be expected in the same logic. This feeding relationship is shown in (7) below.

(7) Underlying representation /ÇjʉÇ/

↓

Voicing assimilation of [j] [Ç̥jʉÇ]

↓

High vowel devoicing [Ç̥j̥ʉ̥Ç]

Another remaining issue is that in my review of literature, there are no studies about whether /i, ʉ/ adjacent to a voiced consonant are in fact never devoiced/deleted or can be devoiced/deleted. Based on my Japanese colleague's and my informal observation, /ʉ/ in words like /sʉ̲mappʉ/ (a Japanese boy band name), in which /ʉ/ is followed by /m/, can be devoiced/deleted at least by some SJ speakers in natural speech. The general assumption that /i, ʉ/ adjacent to a voiced consonant are not devoiced/deleted may need to be reexamined.

3.2.4 Conclusion regarding HVD in [CjʉC] contexts

To conclude, [j] in [CjʉC] could be one of many factors of HVD blocking, and this provides some support to the complex onset /Cj/ hypothesis. However, this may also be due to some other factors, such as infrequent occurrences or orthographic influences. Therefore, the results of my experiment do not strongly support the /Cj/ hypothesis. Meanwhile, the fact that more than 60% of /ʉ/ in /CjʉC/ was devoiced does not reject the /Cj/ hypothesis. Although this experiment is not conclusive, it is still noteworthy that this is the first study of observing HVD patterns in [Cjʉ] morae as well as loanword [Cʉ, Ci] morae.

4 Discussion

4.1 The number of segments and native speaker intuitions

Japanese has a word whose pronunciation is notoriously difficult even for trained native speakers, きゃりーぱみゅぱみゅ /kjaɾi:pámjʉpamjʉ/ 'Kyary Pamyu Pamyu (a Japanese fashion model and singer).' Although the repetition of bilabial onsets makes one of the reasons for the difficulty in articulation, the major difficulty lies in producing the /CjV.CV.CjV.CV/ structure while maintaining the mora-timed rhythm. Native SJ speakers would be able to pronounce /pámʉpamʉ/ much faster than /pámjʉpamjʉ/. This can be explained by analyzing the /mjʉ/ mora as having three segments, which therefore requires a longer time than two or one-segment morae (i.e. /CV/, /V/, and /C/) in articulation. Likewise, in mimetics, only one [Cj] is permitted in one root, and this is called monopalatality: for example, [pjoko-pjoko] 'flip-flap' but not *[pjokjo-pjokjo] (Mester & Itō, 1989). Again, according to native SJ speaker intuitions, this constraint may be in part because a three-segment /CjV/ mora is phonetically so long that consecutive CjV morae greatly disturb the mora-timed rhythm. Such disturbance is not tolerated in SJ, in which the mora has an important role. Compare the three nonsense words written in Japanese romanization (in which the letter <y> corresponds to the glide [j]) in (8):

(8) (a) namanama (b) nyamanyama (c) nyamyanyamya

According to the opinions from some of the participants of the experiment above, native SJ speakers felt that (8, c) with consecutive [CjV] morae was unpronounceable in natural speed, and that (8, b) was pronounceable in natural speed but not as fast as (8, a). Again, my interpretation is that /CjVCjV/ takes longer than /CVCV/ because the former contains six segments while the latter contains four. In contrast, when I asked a few of the Russian participants to pronounce these three with palatalized/soft consonants, they commented that all the three were easily pronounceable in natural speed. Moreover, reportedly one of the Russian participants had studied Japanese for around three months by the experiment, and he commented that according to his impression, the Japanese [Cj]s were more similar to the Russian /Cj/ clusters than the Russian palatalized /C^{j}/s. All this suggests that the syllables in (8, c) behave very differently in Russian and SJ. For Russian speakers, (8, b) and (8, c) are phonologically /n^{j}aman^{j}ama/ and /n^{j}am^{j}an^{j}am^{j}a/ respectively in which /n^{j}/ and /m^{j}/ are single consonants just like /n/ and /m/.

Likewise, SJ consecutive CV morae with palatal consonants are pronounceable in natural speed, as seen in (9).

(9) (a) /tsasatsasa/ (b) /tɕasatɕasa/ (c) /tɕaɕatɕaɕa/

The SJ participants felt that although (9, c) has consecutive palatal (or nonanteriorized) consonants, it was completely pronounceable in natural speed, or that (9, c) might take slightly shorter time than (9, b) due to alternative occurrences of two different places of articulation in (9, b), and shorter time than even (9, a) due to infrequent occurrence of [tsa] in SJ. The reason that all the three nonsense words in (9) are pronounceable in natural speed is because /tsa, sa, tɕa, ɕa/ are all two-segment CV morae. This suggests that consecutive SJ /nja/ or /mja/ morae take longer time not because palatal consonants take longer articulation time, but because /nj/ or /mj/ consist of two segments. Such native- speaker intuitions can support the results of the experiments above.

4.2 Existence of more plausible Japanese [C^{j}]-[C] contrasts

If a language has palatalized consonants, their corresponding plain consonants tend to be velarized for the sake of perceptually clear distinctiveness, as seen in Russian or Irish (Padgett, 2003a). Such contrasts are regarded as [back] contrasts. In this sense, SJ potentially has more plausible [back] contrasts (or [C^{j}]-[C^{ɣ}] contrasts) than the [CjV] versus [CV] contrasts, i.e. the /CiC/ versus /C̥C/ contrasts (phonetically [C^{j}C] vs. [C^{ɣ}C] (or [C^{w}C]), C = voiceless consonant) involved in the aforementioned /i, ɯ/ devoicing/deletion. Historically, in the contrastive palatalized consonants in Russian, for example, allophonically palatalized consonants in Old Russian or Old East Slavic were later phonologized (Padgett, 2003a). More specifically, one of the lax vowels called 'jer', i.e. [ɪ], caused palatalization of the preceding consonant, as in [dɑnʲɪ] 'tribute'. Even after the famous loss of jers in certain positions, the preceding consonants still stayed palatalized, as in [dɑnʲɪ] → [dɑnʲ] (Padgett, 2003a, p. 307). Likewise, in

Irish, contrastive consonant palatalization was conditioned by the following high and mid front vowels before the loss of vowels in final and interior syllables (Greene, 1973). Compared with the history of contrastive palatalization in these languages, the potential palatalized-velarized (or labialized) contrasts in SJ are the following cases; consonants are palatalized by the following [i] or velarized (or labialized) by the following [ʉ], and even where [i] and [ʉ] are phonetically deleted due to the aforementioned High Vowel Devoicing/Deletion (HVD), palatalization and velarization/labialization of the consonants still remain. Examples are shown in (10) and (11) (examples from Vance, 2008, pp. 209-210).

(10) [kʲ] vs. [kˠ/kʷ]: /kiɕo:/ 'weather' vs. /kʉɕo:/ 'bitter smile'
[kʲɕo:] vs. [kˠɕo:] or [kʷɕo:]

(11) [ɕʲ] vs. [ɕˠ/ɕʷ]: /ɕitai/ 'dead body' vs. /ɕʉtai/ 'subject'
[ɕʲtai] vs. [ɕˠtai] or [ɕʷtai]

In the cases of (10) and (11), /kiC/ versus /kʉC/ and /ɕiC/ versus /ɕʉC/ are phonetically [kʲC] versus [kˠC] (or [kʷC]) and [ɕʲC] versus [ɕˠC] (or [ɕʷC]). These cases in SJ behave much like palatalized-velarized consonants in Russian since these contrasts are caused by high vowel deletion compared to [CjV]s with no velarized counterparts. Moreover, unlike the aforementioned examples like /pámʉpamʉ/ versus /pámjʉpamjʉ/, native SJ speakers would be able to pronounce both [kʲɕo:] and [kˠɕo:] (or [kʷɕo:]) equally easily and quickly, and the same is true for [ɕʲtai] and [ɕˠtai] (or [ɕʷtai]) as well. This is because underlyingly both [Cʲ] and [Cˠ(ʷ)] consist of the same number of segments. In this sense, SJ potentially has two completely different contrasting systems: palatalized- velarized (or labialized) contrasts and complex-simplex (Cj-C) contrasts, although consonant sequences like [kʲC] versus [kˠ(ʷ)C] are currently not phonologized in SJ.

5 Limitations and future research

5.1 Limitations about the participants

Because both Japanese-speaking and Russian-speaking participants were recruited in Victoria, Canada, many of them were not monolingual. The Russian speakers' length of residence (LOR) in an English speaking country is ranging from 7 months to 4.5 years (median 2.25 years). The Japanese speakers' LOR is ranging from 1 week to 41 years (median 2.25 years). So there may be some L2 influence on their L1s. However, because the stimuli were written with the orthography in their L1s, I believe that they were in their L1 modes as much as possible during the experiments. Moreover, all the participants reported that they had connections with their L1 speaking friends in Canada. The Japanese participant with LOR 41 years, who may present an extreme case, is married to another Japanese participant with an LOR of 34 years. I believe that their regular

contact with their L1 speakers can help reduce the L2 influence on their L1s. Another limitation is that the number of Russian speakers is not large.

5.2 Complex onsets /Cj/ or rising diphthongs /iV/

Another possible interpretation of the longer duration of [CjV] is that [j] in Japanese [CjV] belongs to the nucleus rather than the onset, that is, [j] is part of a rising (in sonority) diphthong /iV/. The results of this present study cannot judge whether [j] is part of a complex onset or a complex nucleus. This will be examined in future research.

6 Conclusion

I found that Standard Japanese (SJ) [CjV]s, specifically, [gja, nja, mja, bja, ɾja], were highly significantly longer than the [CV] counterparts [ga, na, ma, ba, ɾa], whereas there was no significant difference between the alveo-palatal [dʑa] and the non-palatal [dza]. This indicates that there was no evidence that palatal consonants are longer than their non-palatal counterparts, suggesting that the longer duration of SJ [Cj]s can be better explained by treating [Cj]s as consisting of two segments /CC/s rather than as palatalized consonants /C^{j}/s. To further examine the duration of palatalized consonants, I compared Russian uncontroversial /C^{j}/s [n^{j}a, m^{j}a, b^{j}a, ɾ^{j}a] with the plain/velarized counterparts [na, ma, ba, ɾa]. In unstressed syllables, there was no significant differences between [C^{j}V]s and the [CV] counterparts in duration. In stressed syllables, the absolute durations of [m^{j}a, b^{j}a, ɾ^{j}a] were significantly longer than the [ma, ba, ɾa] counterparts, but not as much as the SJ unaccented [mja, bja, ɾja] versus [mja, bja, ɾja]. However, in the ratios between the target [C$^{(j)}$V]s and the following [ka/ta]s, only [b^{j}a] was significantly longer than [ba] at the $p<0.05$ level, but notably only by 7%. This is much less noticeable compared to the unaccented SJ [bja], which was significantly longer than [ba] with p=2E-37 by 20%. Therefore, I conclude that in terms of duration, SJ [Cj]s behave like consonant clusters, i.e. /CC/ or /Cj/, as opposed to Russian /C^{j}V/s. At least, there is no evidence that SJ [Cj]s are single palatalized consonants. As for High Vowel Devoicing/Deletion (HVD) patterns, I find that [j] in [Cj̥C] can be one of the factors of HVD blocking, but HVD blocking observed in my experiment may also be because infrequent morae are less frequently devoiced, or two-letter morae are less frequently devoiced. The results do not strongly support the /Cj/ hypothesis. Conversely, the fact that /u̥/ in /Cj̥C/ was still devoiced more than 60% does not reject the /Cj/ hypothesis. Since /L/ in /CL/ (L=liquid) onsets in English are also devoiced, devoicing of /j/ in /Cj̥C/ is also expected. Considering the history of Russian and Iris palatalized-velarized contrasts, a potential palatalized-velarized contrasts in Japanese would be /CiC/ versus /Cu̥C/ ([C^{j}C] versus [C^{ɣ}C] or [C^{w}C]) involving HVD, /kiɕo:/ ([k^{j}ɕo:]) versus /ku̥ɕo:/ ([k^{ɣ}ɕo:] or [k^{w}ɕo:]), for example. This contrast in Japanese would be more similar to Russian palatalized-velarized contrasts than the Japanese [Cj]-[C] contrasts.

Acknowledgements

I am grateful to Dr. Ewa Czaykowska-Higgins for supervising my research on Japanese [Cj], and her suggestion of comparing Japanese and a Slavic language and integrating Japanese linguists' feedback about vowel devoicing. I am also grateful to Drs. Hua Lin and Sonya Bird for their valuable advice on this research. Dr. Hua Lin also proofread and revised parts of this paper. I also thank my colleague Natallia Litvin for her advice regarding Russian phonology and help in making the stimuli. Last but not least, I am very grateful to all the 42 participants of the experiments.

References

Fujimoto, M. & Kiritani, S. (2003). Comparison of Vowel Devoicing for Speakers of Tokyo- and Kinki Dialects. *Journal of the Phonetic Society of Japan, 7*(1), 58-69.

Greene, D. (1973). The growth of palatalization in Irish. *Transactions of the Philological Society* 72(1), 127-136.

Grenon, I. (2005). The status of the sound [z] in Japanese. *Revue Langues et Linguistique, 31*, 64–90.

Hall, T. A. (2000). Typological generalizations concerning secondary palatalization. *Lingua, 110*, 1-25.

Hashimoto, T. (1984). A phonological study of "yō-on" on Japanese. *Journal of the Phonetic Society of Japan 175*(4), 1-2.

Hizume, S. (2003). On 茶 as a Kana Expressing Yo-on in Japanese. *Studies in humanities* 39, 114-103.

Hyman, L. M. (2006). Word-prosodic typology. *Phonology*, *23*(02), 225-257.

Kenstowicz, M. J. (1994). *Phonology in generative grammar*. Oxford: Blackwell.

Kida, A. (1998). Syllable and mora. *Journal of the Phonetic Society of Japan 2*(1), 41-49.

Kochetov, A. (2002). *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization.* New York: Routledge.

Kohno, M. (1998). Mora, syllable, and rhythm – a psycholinguistic study. *Journal of the Phonetic Society of Japan 2*(1), 16-24.

Kondo, M. (2000). Vowel Devoicing and Syllable Structure in Japanese. *Japanese/Korean Linguistics, 9*, 125-138.

Labrune, L. (2012). Questioning the universality of the syllable: evidence from Japanese. *Phonology*, *29*(01), 113-152.

Larson-Hall, J. (2004). Predicting perceptual success with segments: a test of Japanese speakers of Russian. *Second Language Research 20*, 32–76.

Litvin, N. (2014). *An Ultrasound Investigation of Secondary Velarization in Russian* (Master's thesis), University of Victoria, BC.

Matsuzaki, H. (1993). A quantitative study of fluctuation in Japanese loanword-phone notation. *Journal of the Department of Japanese, Tohoku University, 3*. 83-94.

McMahon, A. M. (2002). *An introduction to English phonology* (Vol. 22). Edinburgh: Edinburgh University Press.

Mester, R. A., & Itô, J. (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language*, 258-293.

Nogita, A. (2010). *Examination of the [si] and [ʃi] Confusion by Japanese ESL Learners* (Master's thesis), University of Victoria, BC.

Nogita, A. Yamane, N., & Bird, S. (2013). The Japanese unrounded back vowel /ɯ/ is in fact rounded central/front [ʉ̟ - ʏ]. Ultrafest VI.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, *32*(2), 258-278.

Padgett, J. (2003a). The emergence of contrastive palatalization in Russian. *Optimality Theory and language change 56*, 307-335.

Padgett, J. (2003b). Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory*, *21*(1), 39-87.

Padgett, J. (2008). Glides, vowels, and features. *Lingua 118*, 1937–1955

Parker, S. (2012). Sonority distance vs. sonority dispersion – a typological survey. In Steve Parker, editor. *The Sonority Controversy*. Berlin: De Gruyter, 101-65.

Shinohara, S. (2012). *"Miseru Koe" no Tsukurikata [How to make charming voice]*. Tōkyō, Japan: Kōdansha.

Takayama, M. (2003). Issues in Japanese historical phonology. *Journal of the Phonetic Society of Japan. 7*(1), 35-46

Takayama, M. (2005). Asymmetry of medial approximants j and w in Japanese. *Journal of the Phonetic Society of Japan.9*(1), 30-38

Vance, T. J. (2008). *The sound of Japanese*. New York: Cambridge University Press.

Warner, N., & Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America*, *109*(3), 1144-1156.

Yamaoka, T. (2008). On the learning of the English /æ/ sound by Japanese speaking learners of English: Explanation of its levels of difficulty and formulation of its teaching principles. *The Bulletin of Japanese Curriculum Research and Development, 30*(4), 49-58.

Appendix A

無意味な言葉を含む文を、3回ずつ読んで下さい。

無意味な言葉は、アクセントの位置を必ず2番目に。例えば「たかはし（高橋）」は「か」にアクセント。
ただしアクセントを強調し過ぎず、あくまで自然なスピードで読んで下さい。
例文）私たちはノリマチです。×３回

１．私たちはゴ マ モチです。　　２．私たちはギョ ナ マチです 。

３．私たちはチュカボシです。　　４．私たちはガタバコです。

５．私たちはクタボシです。　　６．私たちはフィタバシです。

７．私たちはヒュタバチです　　８．私たちはラカバコです。

９．私たちはレナモチです。　　１０．私たちはヒタバシです。

１１．私たちはニャカバコです。　１２．私たちはジャカバコです。

１３．私たちはピュタボシです。　１４．私たちはティカボシです。

１５．私たちはチカボシです。　　１６．私たちはリャカバコです。

１７．私たちはテュカバシです。　１８．私たちはマタバコです。

１９．私たちはギャタバコです。　２０．私たちはナカバコです。

２１．私たちはフタバチです。　　２２．私たちはフュタバチです。

２３．私たちはキュタボシです。　２４．私たちはプタボシです。

２５．私たちはビャタバコです。　２６．私たちはモラマチです。

２７．私たちはミャタバコです。　２８．私たちはトゥカバシです。

２９．私たちはシュカボシです。　３０．私たちはバタバコです。

３１．私たちはザ カ バコです。	３２．私たちはニョ サ モチです。
Мы гатабако	Она хотела гатап
Мы накабако	Она хотела натап
Мы матабако	Она хотела матап
Мы батабако	Она хотела батап
Мы ракабако	Она хотела гатап
Мы гятабако	Она хотела гятап
Мы някабако	Она хотела някап
Мы мятабако	Она хотела мятап
Мы бятабако	Она хотела бятап
Мы рякабако	Она хотела рятап
Мы гъятабако	Она хотела гъятап
Мы нъякабако	Она хотела нъякап
Мы мъятабако	Она хотела мъятап
Мы бъятабако	Она хотела бъятап
Мы ръякабако	Она хотела ръятап