THE ROLE OF CONTRAST IN THE DISTRIBUTIONAL RESTRICTIONS ON JAPANESE PALATALIZED CONSONANTS*

YU TANAKA
University of California, Los Angeles

1 Introduction

Phonemes or sequences of phonemes can be subject to distributional restrictions. Certain phonological forms in a language appear freely in some contexts but never, or very rarely, in others. Recent studies have shown that the notion of contrast and perceptual distinctiveness plays an important role in such distributional restrictions (Flemming 1995, Gallagher 2010, Padgett 2001 to name a few). From this point of view, grammar disfavors not the phonological forms that are absent from the inventory but rather perceptually indistinct contrasts that they would give rise to. This paper points out previously unnoticed skewed distributional patterns of Japanese palatalized consonants and claims that they are another instance of restrictions that are grounded on perceptual distinctiveness. The first goal of the study is to give a full analysis of the limited distribution of Japanese palatalized consonants. The second goal is to provide empirical evidence that some, if not all, distributional restrictions should be analyzed in terms of constraints on contrasts.

2 Distribution of Japanese palatalized consonants

2.1 Background: Plain-palatalized contrast

Japanese consonant phonemes can be characterized by a systematic distinction between plain and palatalized forms. The basic consonant inventory of Japanese is illustrated with common phonetic forms in (1) below. Place classification is based on the place of plain forms. Notice that, except for glides, all plain consonants can be paired with their non-plain counterparts. The actual phonetic realizations of these so-called "palatalized" consonants vary. Some of them are true palatal (e.g. [ɲ] and [ç]) or alveolo-palatal sounds (e.g. [tɕ] and [ɕ]), while others are sounds with secondary

*I am grateful to the audience at FAJL 7 and the UCLA Phonetics Seminar, especially Ryan Bennett, Bruce Hayes, Alexei Kochetov, Haruo Kubozono, Armin Mester, Megha Sundara, and Kie Zuraw, for their comments and discussions. I would also like to thank my research assistants, Yejin Cho, Edward Nguyen, Valerie Shih and Karen Tsai, for running the experiment reported in this paper. All remaining errors are mine.

(1) Japanese consonants based on Itô and Mester (2003), Itô and Mester (2006)

		Labial		Coronal				Dorsal		Glottal	
		C	Cʲ	C	Cʲ	C	Cʲ	C	Cʲ	C	Cʲ
Obstruent	voiceless	p	pʲ	t	tɕ	s	ɕ	k	kʲ	h	ç
	voiced	b	bʲ	d	dʑ∼ʑ	z	dʑ∼ʑ	g	gʲ		
Sonorant	nasal	m	mʲ	n	ɲ						
	non-nasal	w		ɾ	ɾʲ		j				

palatalization (e.g. [pʲ] and [kʲ]). Given that they constitute a natural class in Japanese phonology (Mester and Itô, 1989),[1] and for the sake of simplicity and consistency, I will refer to all of them as "palatalized consonants" unless further phonetic details are in question (cf. Hume 1994).

The fact that plain and palatalized consonants are in contrast can be illustrated by minimal pairs such as [koː] 'back of hand' vs. [kʲoː] 'today.' It is also known, however, that the plain-palatalized contrast is a contextually-limited phonological contrast: it is maintained before a back vowel /u/, /o/, or /a/, but not before a front vowel /i/ or /e/. Before /i/, all consonants are palatalized, and thus /Ci/ and /Cʲi/ are neutralized into [Cʲi].[2] Before /e/, on the other hand, consonants are never palatalized, and there is no contrast between /Ce/ and /Cʲe/.[3] The distribution of plain and palatalized consonants is summarized in (2) below. (See Itô and Mester (2006) for a plausible account of this distributional pattern based on enhancement of contrast.)

(2) Distribution of plain and palatalized consonants in Japanese

*Ci	Cʲi	Cɯ	Cʲɯ
Ce	*Cʲe	Co	Cʲo
		Ca	Cʲa

2.2 Restrictions based on place and following vowel

We have seen that the distribution of Japanese palatalized consonants is limited to certain contexts. While the restriction pattern shown above is categorical and well documented in the relevant literature, there is also a previously unnoticed gradient restriction on palatalized consonants, which is the main topic of this paper.

To illustrate the restriction pattern, I conducted a corpus study using a Sino-Japanese dictionary, *Kanjigen* (Todo et al., 2001). A Sino-Japanese dictionary was chosen as a source of the data since palatalized consonants are mostly found in the Sino-Japanese lexicon (Labrune, 2012). I extracted all monomorphemic roots containing palatalized consonants of ten different places[4] occurring in

[1] Traditional Japanese linguistics classifies these sounds as segments that constitute special CV moras called *yō-on* 'crooked sounds' as opposed *choku-on* 'straight sounds,' CV moras that are composed of plain consonants and vowels.

[2] Some phonemes such as /p/ and /k/ do not change their primary place of articulation before /i/. However, given that production of these sounds also involves fronting and raising of the tongue body to the hard palate, I assume that the consonants are still palatalized in the sense that is employed in this paper. Thus, the phonetic forms of /pi/ and /ki/ can be transcribed as [pʲi] and [kʲi] in narrow transcriptions (see Itô and Mester 2003:10).

[3] Recent loans can have unaffricated forms before /i/ (e.g. [paːtʲiː] 'party') and also palatalized (affricated) forms before /e/ (e.g. [tɕekkɯ] 'check'). See Itô and Mester (1995) for more examples and discussions.

[4] In the Sino-Japanese stratum, the sound [p] occurs (arguably) only as an allophone of /h/ in compound formation (e.g. /kaɴ-hai/ → [kam**p**ai] 'toast'). Thus, /p/ was excluded from the search. /dʲ/ was collapsed with /zʲ/ since the two phonemes are neutralized, both being variably realized as [dʑ] or [ʑ].

five vowel contexts where plain and palatalized forms are contrastive: /a/, /o/, /oː/, /ɯ/ and /ɯː/. Then, the frequency of each consonant in each vowel context was counted.

One note about the frequency count is in order. Entries in a Sino-Japanese dictionary are monomorphemic roots written with one kanji or Chinese character. Due to the fact that words were borrowed from different regional languages of China at different times, one same kanji can have multiple readings.[5] For example, the root for 'woman' can be read [(d)ʑo], [ɲo], and less commonly [ɲoː], depending on the context. In some cases, one reading is more common than the others, but in other cases, all readings seem more or less equally common. For this reason, I included all possible readings in the count. Thus, the root for 'woman' was counted three times.

The database contained 2,335 roots, including overlapping entries. From the observed number, the expected number of each palatalized consonant occurring in each vowel context was calculated using the method employed in Kawahara et al. (2006).[6] The table in (3) shows the number of roots sorted by consonant places and vowel contexts. Expected numbers are shown in parentheses.

(3) Number of roots containing palatalized consonants

		a	o	oː	**ɯ**	**ɯː**
Sibilant	/sʲ/ [ɕ]	124 (115)	107 (90)	267 (374)	**205** (106)	**99** (116)
fricatives	/zʲ/ [ʑ]∼[dʑ]	59 (60)	44 (47)	161 (195)	**95** (55)	**61** (60)
Others	/tʲ/ [tɕ]	23 (32)	30 (26)	117 (106)	**7** (30)	**50** (32)
	/dʲ/ [ʑ]∼[dʑ]	— Collapsed with /zʲ/ —				
	/kʲ/ [kʲ]	38 (47)	41 (37)	185 (152)	**0** (43)	**62** (47)
	/gʲ/ [gʲ]	14 (11)	13 (12)	73 (48)	**0** (13)	**2** (12)
	/nʲ/ [ɲ]	8 (12)	6 (10)	47 (40)	**0** (11)	**24** (12)
	/ɾʲ/ [ɾʲ]	16 (21)	19 (17)	77 (68)	**1** (19)	**33** (21)
	/hʲ/ [ç]	27 (14)	4 (11)	66 (47)	**1** (13)	**3** (15)
Labials	/bʲ/ [bʲ]	7 (7)	0 (7)	53 (29)	**0** (8)	**2** (9)
	/mʲ/ [mʲ]	19 (9)	0 (7)	43 (30)	**0** (8)	**2** (9)

The main focus here is on the two high back vowel contexts /ɯ/ and /ɯː/, which are shown in boldface. Generally speaking, palatalized consonants before short /ɯ/ are much rarer than those before long /ɯː/. For eight out of ten consonants, the number of roots containing the form /Cʲɯ/ is smaller than the expected number, indicating that they are underrepresented. In fact, for places such as /kʲ/, /gʲ/, /nʲ/, /bʲ/ and /mʲ/, there is no root containing /Cʲɯ/ at all. The sibilant fricatives /sʲ/ and /zʲ/ constitute a class of exceptions, appearing freely before short /ɯ/.

Additionally, the palatalized labials /bʲ/ and /mʲ/ appear to be restricted not only before short /ɯ/ but also before long /ɯː/.[7] It is worth noting here that the two roots containing a /bʲɯː/ sequence and the two roots containing a /mʲɯː/ sequence are indeed the same roots. The root

[5] Sino-Japanese readings, or *on-yomi*, are generally classified into four types according to their region and time of origin. Kanjis can also have Japanese readings, or *kun-yomi*, which are based on the pronunciation of native words that have the same or similar meanings. Japanese readings were not considered in the corpus study presented here.

[6] For example, in the table in (3), the total number of roots containing /sʲ/ is 804 (the sum of the /sʲ/ row) and the number of roots containing /a/ is 335 (the sum of the /a/ column). Since there are 2,335 roots in total, the probability of /sʲ/ is 802/2,335 ≈ 0.3435, and the probability of /a/ is 335/2,335 ≈ 0.1435. The probability of both events occurring is 0.3435*0.1435 ≈ 0.0493. Thus, the expected number of roots containing /sʲa/ is 0.0493*2,335 ≈ 115.

[7] We also notice the scarcity of /hʲɯː/. This could be due to the fact that /h/ was historically a labial sound.

/bʲɯː/ meaning 'error' and the root /bʲɯː/ meaning 'to bind, to entangle' both have another reading /mʲɯː/. Since all possible readings were included in the root count as stated above, these roots were counted twice.[8]

To summarize, the data indicate the following. First, there is a general distributional restriction on /Cʲɯ/ but not on /Cʲɯː/. That is, the distribution of palatalized consonants is conditioned by the length of the following vowel /ɯ/. I will refer to this as *vowel length-based restriction*. Second, palatalized sibilant fricatives (/Sʲ/) are exempt from such restrictions, while palatalized labials (/Bʲ/) are restricted even when followed by long /ɯː/. That is, the degree of restriction on palatalized consonants varies depending on their place. I will refer to this as *place-based restriction*. The restriction patterns are summarized in the table in (4) below. In the remainder of the paper, I investigate the basis of these two kinds of distributional restrictions.

(4) Summarized restriction patterns

	ɯ	ɯː
Sibilant fricative	✓Sʲɯ	✓Sʲɯː
General	* Cʲɯ	✓Cʲɯː
Labial	* Bʲɯ	* Bʲɯː

3 The basis of the restrictions

I have shown that Japanese palatalized consonants are subject to place-based and vowel length-based restrictions. In this section, I first outline possible accounts and claim that one based on the perceptual distinctiveness of contrast can give a unified account of the two kinds of restrictions.

3.1 A diachronic account

One might attribute the vowel length-based restriction to diachronic factors. Since many palatalized consonants were formed via a historical sound change */Ciɯ/ > /Cʲɯː/ involving compensatory lengthening (Labrune 2012), the rarity of /Cʲɯ/ could be a mere lexical gap. In other words, palatalized consonants are not found before short /ɯ/ because it is not an environment from which they were historically derived. From this viewpoint, the synchronic grammar of Japanese plays no role in the distributional skew.

This diachronic account, however, has several weaknesses. First, it does not explain why the distribution of palatalized consonants are also conditioned by their place. As we have seen, palatalized sibilant fricatives appear freely before short /ɯ/, which implies that the historical sound change */Ciɯ/ > /Cʲɯː/ is not the only source of palatalized consonants. Palatalized labials, on the other hand, are virtually absent even before long /ɯː/. A diachronic account would not be able to capture this difference based on the place of consonants.

Second, data from loanword adaptation and casual speech suggest that there is some synchronic grammatical force that restricts palatalized consonants before short /ɯ/. Although /Cʲɯ/ is generally allowed in borrowings and used to adapt consonant-glide sequences in English,

[8] Interestingly, the dictionary lists examples of compounds where these roots are read [bʲɯː], such as *gobyū* [gobʲɯː] 'error, mistake,' but it does not give any examples where they are read [mʲɯː]. This suggests that the sequence /mʲɯː/ is extremely rare. The scarcity of /mʲɯ(ː)/ was noted by Kindaichi (1957). He states that the only Japanese words that contain /mʲɯ(ː)/ are two family names, *Mamyūda* [mamʲɯːda] and *Ōmamyūda* [oːmamʲɯːda].

substitution of /Cʲɯ/ with /Ci/ is common in some lexical items (e.g. *simulation* |sɪ**mjə**leɪʃən| → [ɕi**mʲɯ**ɾeːɕoɴ]∼[ɕi**mi**ɾeːɕoɴ], or also possibly [**ɕɯmi**ɾeːɕoɴ] with metathesis, and *community college* |kə**mju**nəti kɑlɪdʒ| → [ko**mi**kaɾe] with compound truncation, and not *?[ko**mʲɯ**kaɾe]). In casual or rapid speech, even palatalized sibilant fricatives followed by /ɯ/, that look otherwise exempt from the restriction as we have seen, can undergo the same substitution process (e.g [**ɕɯ**ʑɯtsɯ]∼[**ɕi**ʑɯtsɯ] 'surgery,' see Labrune 2012 for more data). These data show that the form /Cʲɯ/ is avoided in the synchronic grammar and that the skewed distribution of palatalized consonants is not just a vestige of historical sound change.

3.2 An account based on constraints on individual forms

We have seen that the restrictions on Japanese palatalized consonants need to be explained in terms of a synchronic grammar. As a first attempt, let us propose Markedness constraints that penalize the forms that are marginal in the inventory. We can think of constraints such as *Cʲɯ, which bans a sequence of a palatalized consonant and a short [ɯ], and *Bʲɯː, which bans a sequence of a palatalized labial and a long [ɯː]. Our job is then to seek out motivations for these constraints.

One can claim that the constraint *Cʲɯ is grounded on ease of articulation. Moving the tongue body from front to back between adjacent segments requires some articulatory effort (see FRONT-BACK proposed in Flemming 1995). *Cʲɯ bans such effortful sequences. Although the motivation appears to be solid, the situation at hand is more complicated. Recall that palatalized consonants appear freely before long /ɯː/. If any tongue movement from front to back between adjacent segments involves articulatory effort, the sequence /Cʲɯː/ should be as bad as /Cʲɯ/. The situation looks even more puzzling when we consider the place-based restriction. If the constraint *Bʲɯː has the same articulatory grounding as *Cʲɯ, we need to claim that a sequence of a palatalized labial and a long high back vowel is more effortful than other /Cʲɯː/ forms which are not restricted. To the best of my knowledge, there is no articulatory reason why /Bʲɯː/ forms should be specifically disfavored. Thus, articulatory constraints on particular sequences would not suffice to explain essentially why there are two kinds of restrictions.

3.3 An account based on constraints on contrasts

I have shown that an explanation based on diachronic factors and an analysis based on constraints on particular sound sequences are both insufficient to account for the distributional restrictions on palatalized consonants in Japanese. As an alternative account, I propose an analysis based on constraints on contrasts (Flemming, 1995) arguing that the restrictions are driven by the need to eliminate perceptually weak contrasts such as /kʲɯ/ vs. /ki/ and /bʲɯː/ vs. /biː/.

3.3.1 A contrast-based account of *Cʲɯ and ✓Cʲɯː

The production of palatalized consonants involves a primary articulation with simultaneous raising and fronting of the tongue body to the hard palate. Kochetov (2004) shows that the secondary palatalization gesture can overlap with the production of adjacent vowels.[9] Note that this overlapping gesture could potentially affect the perception of following vowels, especially in the case of a short /ɯ/. Homma (1985/1992) reports that the phonetic realization of Japanese /ɯ/ is considerably affected by its adjacent consonant and transcribes it as [ɯ̈] or [ɨ] after /t/, /s/,

[9] Kochetov also shows that the timing of the gestures differs depending on the language, place and syllable position.

/n/, and /j/. When /ɯ/ is overlapped by a preceding palatalized consonant, not only does tongue fronting raise the F2 of the vowel, but the formant transition also shortens the steady state of the formants which characterize the vowel's quality. This can result in misperception of /ɯ/ as /i/, and hence neutralization of /C^{j}ɯ/ to /Ci/. I propose that there is a constraint on perceptually indistinct contrasts and that /C^{j}ɯ/ is scarce in the lexicon due to its perceptual similarity to /Ci/.

With this contrast-based analysis, the fact that palatalized consonants are not restricted before long /ɯː/ can be explained as follows. If the following vowel is long, there is less of a chance of misperception. Tongue fronting especially affects the vowel quality of the portion that is adjacent to the consonant, and the portion after the formant transition reaches its end is less affected. This means that the steady state of long /ɯː/ has a relatively long duration, which can serve as a stable perceptual cue. This helps maintain the contrast between /C^{j}ɯː/ and /Ci(ː)/. I propose that the vowel length-based restriction is driven by this difference in perceptual distinctiveness.

Recall now the data from loanword adaptation and casual speech (§3.1). A /C^{j}ɯ/ sequence is always repaired by changing it to /Ci/. These data also indicate perceptual similarity between /C^{j}ɯ/ and /Ci/. Such repair never happens with /C^{j}ɯː/.

3.3.2 A contrast-based account of *B^{j}ɯ and *B^{j}ɯː

The contrast-based analysis also explains why palatalized labials are restricted not only before /ɯ/ but also before /ɯː/. Cross-linguistically, palatalization on labials tends to be realized as a palatal off-glide (Bhat 1978). Kochetov (2004) argues that this delayed palatalization for labials is motivated by perceptual enhancement. It is hard for listeners to perceive a palatal gesture articulated simultaneously with a labial because the lip closure masks the acoustic effects of the tongue movement. Thus, palatalization is delayed in order to enhance the contrast between palatalized labials and plain labials (e.g. /b^{j}a/ [bja] vs. /ba/ [ba]).

Note, however, that this can bring about yet another risk of neutralization. As discussed above, an overlapping palatalization gesture affects the quality of the following vowel. Delayed palatalization yields further F2 raising and shortening of the steady state of the vowel formants. This makes the identification of the following /ɯ/ harder than other C^j__ contexts, even when the vowel is long. By excluding /B^{j}ɯː/ from the inventory, a language can eliminate the weak contrasts /Bɯː/ vs. /B^{j}ɯː/ and /B^{j}ɯː/ vs. /Biː/, and maintain a strong contrast between /Bɯː/ and /Biː/. This is illustrated by a schematic image of possible labial-vowel inventories in (5).

(5) Possible inventories of labial-vowel sequences

a. Three-way contrast

bɯː $\xleftrightarrow[\textit{weak}]{\quad}$ b^{j}ɯː $\longleftrightarrow$ biː

b. Enhancement of /Bɯː/-/B^{j}ɯː/ by delayed palatalization

bɯː $\xleftrightarrow[\textit{enhanced}]{\quad}$ bjɯː $\xleftrightarrow[\textit{weakened}]{\quad}$ biː

c. Two-way, elimination of /B^{j}ɯː/

bɯː $\xleftrightarrow[\textit{contrast strong}]{\cancel{\text{*bjɯː}}}$ biː

Thus, I propose that /B^j/ is restricted not only before /ɯ/ but also before /ɯː/ since /B^{j}ɯː/ cannot maintain its perceptual distinctiveness from /Bi(ː)/ due to the delay of palatalization.

3.3.3 A contrast-based account of ✓Sʲɯ

Why sibilant fricatives are exempt from the restrictions is explained as follows. First, Japanese /sʲ/ and /zʲ/ undergo full-palatalization and are realized as alveolo-palatal [ɕ] and [ʑ], which involve less overlapping gestures with the following vowel compared to segments with secondary palatalization. Second, sibilant fricatives are reported to contain not only strong cue to their own place (Wright, 2004) but also strong acoustic information about following vowels, especially the F2 (Soli, 1981). This helps maintain the contrast between /Sʲɯ/ and /Si/, and thus palatalized sibilant fricatives can exceptionally stand before short /ɯ/ (except in casual or rapid speech.)

3.4 Summary

I have proposed that place-based restriction and vowel length-based restriction are both grounded on perceptual distinctiveness. There are constraints that penalize perceptually indistinct contrasts, and the forms /Cʲɯ/ and /Bʲɯː/ are restricted in the lexicon because the contrasts /Cʲɯ/ vs. /Ci/ and /Bʲɯː/ vs. /Biː/ are indistinct. /Sʲɯ/ is exceptionally allowed because the contrast /Sʲɯ/ vs. /Si/ is sufficiently distinct.

4 Experiment

To test the hypotheses, I conducted a gating experiment. Native English speakers were presented with part of monosyllabic Japanese words having the structures /Ciː/ and /Cʲɯː/. At each gate, they received a longer portion of the word and were asked to judge whether the following vowel would sound like English /i/ or /u/. The contrast-based analysis makes the following predictions. (i) In general, the vowel in /Cʲɯː/ will be identified as /i/ at earlier gates because of the perceptual similarity between /Cʲɯ/ and /Ci/. (ii) The vowel in /Bʲɯː/ will be identified as /i/ for longer gates compared to other /Cʲɯː/ forms because of delayed palatalization. (iii) The vowel in /Sʲɯː/ will be identified as /u/ at earlier gates because of the perceptual advantage of sibilant fricatives.

4.1 Methods

4.1.1 Stimuli

Monosyllabic Japanese words having the structures /Ciː/ and /Cʲɯː/ with twelve different consonants were used as the stimuli. The list of the stimuli is given in (6) below.[10]

(6) Experiment stimuli

	Cʲɯː	Ciː	*feature*		Cʲɯː	Ciː	*feature*
/s/	[ɕɯː]	[ɕiː]	[+str, +cont]	/ɾ/	[ɾʲɯː]	[ɾʲiː]	[–str, –lab, +cont]
/t/	[tɕɯː]	[tɕiː]	[+str, –cont]	/h/	[çɯː]	[çiː]	
/d/	[dʑɯː]	[dʑiː]		/p/	[pʲɯː]	[pʲiː]	[+lab, –cont]
/k/	[kʲɯː]	[kʲiː]	[–str, –lab, –cont]	/b/	[bʲɯː]	[bʲiː]	
/g/	[gʲɯː]	[gʲiː]		/m/	[mʲɯː]	[mʲiː]	
/n/	[ɲɯː]	[ɲiː]		/ɸ/	[ɸʲɯː]	[ɸʲiː]	[+lab, +cont]

[10] /zʲ/ is not included as it is neutralized with /dʲ/ and in the recording of the stimuli (see below) all speakers produced the phoneme as [dʑ] with affrication, a common allophone in word-initial position. /ɸʲ/ is a phoneme used in loanwords and was included to test whether the hypothesis about labial sounds applies to this new phoneme.

The words are all unnaccented. Except for one existing content word [kʲɯː] 'sudden,' they are nonsense or onomatopoeic words. Additionally, [tʲiː], [tʲɯː], [dʲiː] and [dʲɯː], sequences that are found only in loanwords, were included as the stimuli for practice sessions. To make audio stimuli, the words were pronounced by four female native speakers of Tokyo Japanese who were naive to the purpose of the experiment. The recording was done in a sound-attenuated recording booth in the UCLA Phonetics Lab. Their speech was recorded through a Shure head-mounted SM10A microphone into a computer sound card via a pre-amp in stereo at the sampling rate of 44,100 kHz with 32 bit resolution. Each word was embedded in a carrier sentence: *Haha-wa _____-to itta.* 'Mother said _____.' The sentences were presented one by one on a computer screen and the speakers were asked to read them aloud. The target words were written in *katakana* orthography.

The target words were cut out and then edited using Praat (Boersma and Weenink, 2014). 150ms before the beginning of the vowel (i.e., the beginning of periodic cycles on the waveform) were cut out and saved to be used as a first gate. Then, 15ms of the vowel portion were successively added to create each gate. Thus, the first gate was composed of 150ms of the consonant part of a /Ciː/ or /Cʲɯː/ word (including silence before the consonant). The second gate was composed of 150ms of the consonant part plus 15ms of the vowel portion. The third gate was the consonant part with 30ms of the vowel portion. Fifteen gates were created for each stimulus word in this manner.

4.1.2 Participants

Fifteen native speakers of English participated in the experiment. The participants were all undergraduate students at the University of California, Los Angeles, and were recruited via the UCLA Psychology Department Subject Pool. They received one course credit for participation.

4.1.3 Procedure

The experiment was designed and run in Praat. In the instruction session, participants were told that they would be listening to part of non-English words and judge whether the following vowel would sound closer to the English vowel /i/ as in the word *tea*, or /u/ as in the word *due*. The test started with a practice session. Participants were auditorily presented with fifteen gates of [tʲiː], [tʲɯː], [dʲiː] and [dʲɯː]. The fifteen gates of each stimulus word were grouped as a block. Each block started with the first gate of the stimulus, and then moved on to the next gate until the fifteenth gate. At each gate, they were asked to judge whether the following vowel was closer to English /i/ or /u/. To make their selection, they could click on either a button marked *ee* or one marked *u*.[11] They also gave a confidence rating on a scale of 1 to 100 for each response. In the main session, they were presented with fifteen gates of /Ciː/ and /Cʲɯː/ words with twelve different consonants. The task consisted of twenty-four stimulus blocks, whose order was randomized for every participant. They were given a short break every six blocks.

4.2 Results

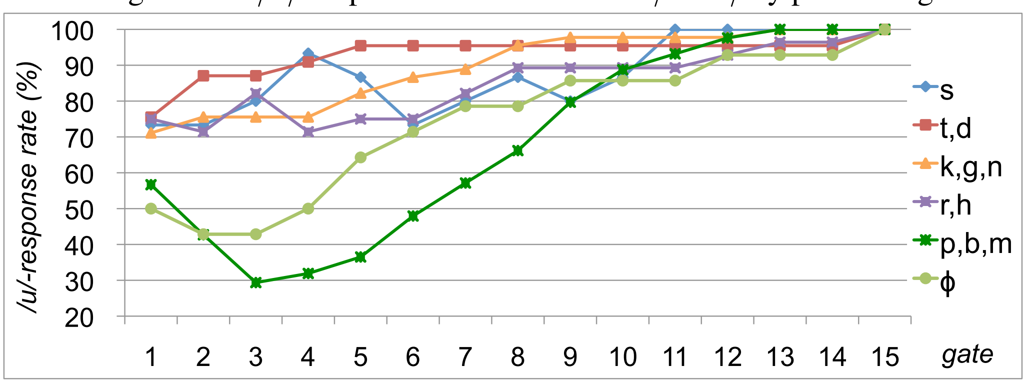
The responses were converted to numerical values. For identification of the vowel in /Ciː/-type stimuli, /i/-responses were counted as the "correct" answers and converted into 1's and

[11] The spelling *u* was used instead of *oo* because of the following reason. To English speakers, a sequence of a palatalized consonant and a vowel can sound like a consonant-glide-vowel sequence. In the English spelling system, *u* is most commonly used to represent /u/ after a consonant-glide sequence (e.g. *cute*, *mule*, etc.).

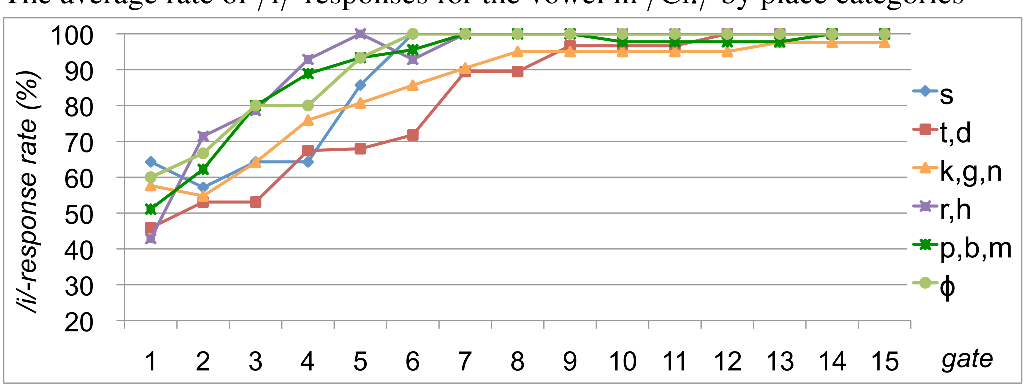
/u/-responses were counted as "incorrect" answers and converted into 0's. For identification of the vowel in /Cʲɯː/-type stimuli, /u/-responses were converted into 1's as the "correct" answers and /i/-responses were converted into 0's as "incorrect" answers. Some participants did not provide the correct answer for some stimuli even at the last gate. Since this indicates the possibility that the Japanese vowels /i/ and /ɯ/ were not perceptually mapped to the intended English vowels /i/ and /u/, possibly due to acoustic differences between the vowels in the two languages, these responses were excluded. Out of the total 180 data points, 15 were excluded for this reason.

The results appear in (7). (7a) plots the average rate of /u/-responses in identification of the vowel in /Cʲɯː/-type stimuli, sorted by consonant classification given in (6). (7b) plots the average rate of /i/-responses in identification of the vowel in /Ciː/-type stimuli. In both plots, the *y*-axis represents the rate of responses in favor of a specific vowel (either /u/ or /i/) while the *x*-axis represents gates. As a general tendency, participants made more incorrect responses at earlier gates and correct responses at later gates. This indicates that /Cʲɯː/ and /Ciː/ are more confusable when the vowel is shorter. In (7a), it can be seen that vowels in /pʲɯː/, /bʲɯː/, /mʲɯː/ and /ɸʲɯː/ were identified as /i/ by many subjects at early gates (the bottom green and light green lines) compared to other places, indicating that /Bʲɯː/ is particularly confusable with /Biː/. The results do not show clear effects of the perceptual advantage of sibilant fricatives (the blue line).

(7) a. The average rate of /u/-responses for the vowel in /Cʲɯː/ by place categories



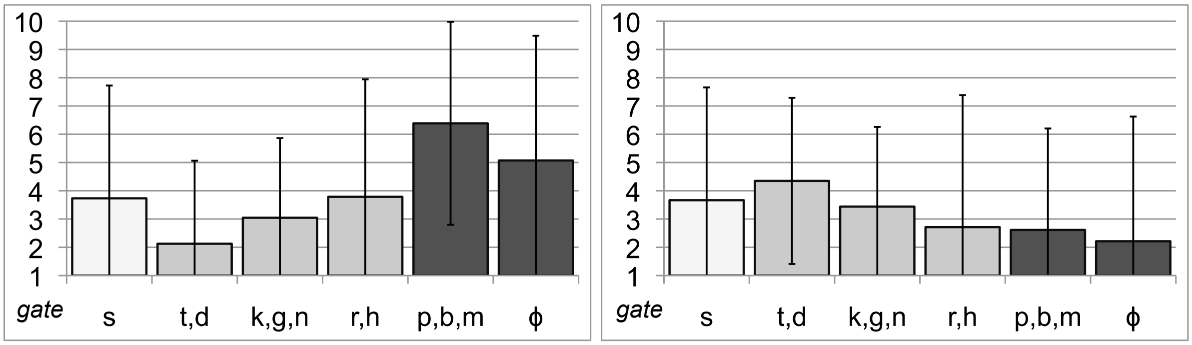
b. The average rate of /i/-responses for the vowel in /Ciː/ by place categories



For the dependent variable of a statistical analysis, each subject's isolation point for each stimulus word was calculated. Isolation points refer to the number of gates listeners needed to provide the correct response without making any further changes thereafter (Grosjean, 1996).[12] (8a) graphs the average isolation points for identification of the vowel in /Cʲɯː/-type stimuli by place categories. (8b) graphs the average isolation points for vowel identification in /Ciː/-type stimuli. Error bars represent plus/minus one standard deviation.

(8) The average isolation points by place categories (error bars: ±1 SD)

a. /Cʲɯː/ stimuli b. /Ciː/ stimuli



To compare isolation points in different conditions (place categories) in each stimulus type (/Cʲɯː/ and /Ciː/), a general linear mixed-effects model (Baayen, 2008) was constructed using the `lmer()` function of the `lme4` package (Bates et al., 2011) implemented in R (R Development Core Team, 1993-2014) with identification points as a dependent variable and conditions as an independent variable. Subjects, stimulus items and speakers of the stimuli were entered as random factors. Post-hoc pair-wise comparisons with Tukey's adjustment were conducted using the functions `glht()` and `mcp( = "Tukey")` in the `multcomp` package (Hothorn et al., 2008).

The analysis revealed the following. In vowel identification in /Ciː/ words, none of the differences between the conditions was significant. In vowel identification in /Cʲɯː/ words, the average identification point for /pʲɯː/, /bʲɯː/ and /mʲɯː/ (6.39 gates) was significantly higher than that for /kʲɯː/, /gʲɯː/ and /nʲɯː/ (3.04 gates), and that for /tʲɯː/ and /dʲɯː/ (2.13 gates) (*p*s < 0.01). The differences between the /pʲɯː/-/bʲɯː/-/mʲɯː/ condition and any other conditions, /sʲɯː/, /rʲɯː/-/hʲɯː/ and /ɸʲɯː/, were not significant ($p = 0.28$, $p = 0.17$, $p = 0.79$, respectively). This indicates that, although the effects are not as clear-cut as was predicted, the contrast between /Bʲɯː/ and /Biː/ is generally less distinct compared to other /Cʲɯː/-/Ciː/ contrasts. Also, the differences between the identification point in the /sʲɯ/ condition (3.73 gates) and that in other conditions did not turn out to be significant. That is, the effects of the perceptual advantage of sibilant fricatives were not found in this experiment. The next section will discuss these results.

4.3 Discussions

We have seen that the results of the experiment only partially supported the hypothesis about the perceptual similarity between /Bʲɯː/ and /Biː/. Note, however, that the average identification point for /pʲɯː/, /bʲɯː/ and /mʲɯː/ is numerically the highest and that for /ɸʲɯː/ is the second highest. That is, the general trend still suggests that palatalized labials make the identification of

[12] Another common dependent variable in gating experiments is recognition points, which refer to the point where listeners give the correct response without further changes and give an 80% or higher confidence rate to their response. Due to a problem in the experiment implementation, some participants did not provide reliable confidence rates, and thus recognition points were not put into analysis in this paper.

the following vowel harder compared to other places. Lack of statistical power due to the small number of participants (n = 15) made it difficult to make a decisive conclusion about whether these trends are robust effects. Further testing with more participants would solve the issue.

Secondly, the results of this experiment did not provide support for the hypothesis about the strong perceptual distinctiveness between /Sʲɯ/ and /Si/. To interpret this result, it seems necessary to take the effects of the stimulus producers into consideration. Recall that some data points were excluded because the vowels were not mapped to the intended English vowels by participants even at the last gate. It is worth noting that out of the fifteen discarded data points, eleven were recordings made by two speakers (with five produced by one speaker and six produced by the other speaker). This implies that the vowel quality of these speakers did not fall in the typical acoustic range of the two English vowels /i/ and /u/ and made the task particularly difficult.

The effects of one of these speakers look particularly prominent in vowel identification in /sʲɯː/ words. Even though no data were discarded for /sʲɯː/ since all participants eventually identified the vowel as /u/ at the fifteenth gate, the identification points for the stimuli produced by one speaker is much higher compared to those for the stimuli produced by other speakers.[13] For these stimuli, participants consistently changed their answers until the identification point, which is a tendency that cannot be seen in answers to other stimuli. Thus, it is possible that participants had difficulty identifying the vowel in these stimuli not because of the consonantal context but simply because of the acoustic properties of this speaker's vowel /ɯ/. Again, the small number of participants made it difficult to conduct further analyses of the speakers' effects. It is hoped that further testing with some change in the experimental design (e.g. exclusion of the stimulus speaker in question, using other languages with palatalized consonants for stimuli, or using speakers of other languages as participants) will clarify the problem.

5 Conclusion

This paper has pointed out previously unnoticed restrictions on palatalized consonants in Japanese. With corpus data, I have shown that their distribution is conditioned by their place and the length of the following vowel. I have proposed an analysis based on contrasts, which gives a unified account of the two kinds of restrictions. I have also presented a gating experiment which was designed to test the proposal. The results gave partial support to the hypothesis. Further investigation awaits to settle the remaining issues, but the present study makes a small contribution to the field of phonology by providing another instance of distributional restrictions that are grounded on the perceptual distinctiveness of contrast.

References

Baayen, Harald R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Bates, Douglas, Martin Maechler, and Ben Bolker. 2011. lme4: Linear mixed-effects models using S4 classes. R package.

[13] Once the data points with this speaker are removed (four out of fifteen data points), the average identification point for /sʲɯː/ goes down to 1.55 gates (SD: 1.21) from 3.73 gates (SD: 4). In nine out of the remaining eleven data points, identification points are all 1's, meaning that participants identified the vowel as /u/ as early as the first gate. This indicates that the results with one stimulus producer considerably raised the average identification point.

Boersma, Paul, and David Weenink. 2014. Praat: Doing phonetics by computer [Computer program]. version 5.3.75. *http://www.praat.org/* .

Flemming, Edward. 1995. Auditory representations in phonology. Doctoral Dissertation, University of California, Los Angeles.

Gallagher, Gillian. 2010. The perceptual basis of long-distance laryngeal restrictions. Doctoral Dissertation, Massachusetts Institute of Technology.

Grosjean, François. 1996. Gating. *Language and the Cognitive Processes* 11:597–604.

Homma, Yayoi. 1985/1992. *Nichi-Eigo no onkyouonseigaku. [Acoustic phonetics in English and Japanese].* Kyoto: Yamaguchi Syoten.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50:346–363.

Hume, Elizabeth. 1994. *Front vowels, coronal consonants and their interaction in nonlinear phonology*. New York: Garland.

Itô, Junko, and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In *University of Massachusetts occasional papers in linguistics [UMOP] 18: Papers in Optimality Theory*, ed. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, 181–209. Amherst: GLSA.

Itô, Junko, and Armin Mester. 2003. *Japanese morphophonemics*. Cambridge, MA: MIT Press.

Itô, Junko, and Armin Mester. 2006. Systemic markedness and faithfulness. In *Proceedings of the 39th Meeting of the Chicago Linguistics Society*, ed. Jonathan Cihlar, Franklin Amy, David Kaiser, and Irene Kimbara, 665–689. Chicago: Chicago Linguistics Society.

Kawahara, Shigeto, Hajime Ono, and Kiyoshi Sudo. 2006. Consonant co-occurrence restrictions in Yamato Japanese. In *Japanese/Korean Linguistics*, Vol. 14, 27–38. Stanford: CSLI Publications.

Kindaichi, Haruhiko. 1957. *Nihongo [Japanese].* Iwanami Shinsho. Tokyo: Iwanami Shoten.

Kochetov, Alexei. 2004. Perception of place and secondary articulation contrasts in different syllable positions: Language-particular and language-independent asymmetries. *Language and speech* 47:351–382.

Labrune, Laurence. 2012. *The Phonology of Japanese (The Phonology of the World's Languages)*. Oxford: Oxford University Press.

Mester, Armin, and Junko Itô. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65:258–293.

Padgett, Jaye. 2001. Contrast dispersion and Russian palatalization. In *The role of speech perception in phonology*, ed. Elizabeth Hume and Keith Johnson, 187–218. New York: Academic Press.

R Development Core Team. 1993-2014. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Soli, Sigfrid D. 1981. Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *Journal of the Acoustical Society of America* 70:976–984.

Todo, Akiyasu, Akira Matsumoto, Akira Takeda, and Yoshimitsu Kano, ed. 2001. *Kanjigen*. Tokyo: Gakushu Kenkyu-sha.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. In *Phonetically-based phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade, 34–57. Cambridge: Cambridge University Press.