

## **UC Santa Cruz**

### **UC Santa Cruz Electronic Theses and Dissertations**

#### **Title**

Degemination in Japanese Loanwords from Italian

#### **Permalink**

<https://escholarship.org/uc/item/8p95v6sw>

#### **Author**

Morimoto, Maho

#### **Publication Date**

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

**DEGEMINATION IN JAPANESE LOANWORDS FROM ITALIAN**

A thesis submitted in partial satisfaction  
of the requirements for the degree of

MASTER OF ARTS

in

LINGUISTICS

by

**Maho Morimoto**

September 2015

The thesis of Maho Morimoto  
is approved:

---

Professor Armin Mester, Chair

---

Professor Junko Itô

---

Associate Professor Grant McGuire

---

Tyrus Miller  
Vice Provost and Dean of Graduate Studies



# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>3</b>
2.1	Italian and Japanese Phoneme Inventories . . . . .	3
2.2	The Status of Geminate in Italian and Japanese . . . . .	5
2.2.1	Geminate in Italian . . . . .	5
2.2.2	Geminate in Japanese . . . . .	6
2.2.3	Geminate in Orthographic Systems . . . . .	8
2.3	Phonological Representation of Geminate . . . . .	9
<b>3</b>	<b>The Basic Data and the Proposal</b>	<b>11</b>
3.1	The Data . . . . .	13
3.2	The Proposal . . . . .	17
<b>4</b>	<b>Theoretical Assumptions and the Basic Analysis</b>	<b>22</b>
4.1	Output-output Faithfulness . . . . .	22
4.2	Positional Faithfulness . . . . .	23
4.3	Deriving the positional effect in zuccotto . . . . .	25
<b>5</b>	<b>Variation</b>	<b>27</b>
5.1	Free Variation . . . . .	27
5.2	Compensatory Lengthening . . . . .	30
5.3	Implicational hierarchy of faithfulness . . . . .	32
5.4	Summary and Predictions . . . . .	36
<b>6</b>	<b>A Nonce-adaptation Survey</b>	<b>38</b>
6.1	Methods . . . . .	38
6.1.1	Participants . . . . .	38
6.1.2	Material . . . . .	39
6.1.3	Procedure . . . . .	42
6.2	Results and Discussion . . . . .	43
6.2.1	Effect of Types of Geminate . . . . .	44
6.2.2	Effect of Position of Geminate . . . . .	46
6.2.3	Interaction of Position and Types of Geminate . . . . .	51
6.3	Summary and Further Development . . . . .	54
<b>7</b>	<b>Conclusion</b>	<b>55</b>
	<b>Appendix The Survey</b>	<b>57</b>

## List of Figures

1	Genres of Italian borrowings in Japanese . . . . .	13
2	Frequency of rating responses . . . . .	43
3	Average ratings for each type of input . . . . .	44
4	Average ratings for outputs containing liquid geminates (“yes”) and outputs without one (“no”). . . . .	45
5	Preferences for operation per types of geminates. Note that the same rating is used twice to calculate this, once for weak geminates, once again for strong geminates. . . . .	46
6	Average Rating for outputs including geminates in either weak position (C1) or strong position (C2) (outputs without any geminates (i.e. no preservation) were eliminated from the analysis). . . . .	47
7	Average ratings for outputs depending on whether they contained a geminate, (a) in the weak position or (b) strong position. . . . .	47
8	Preferred operation for geminates in (a) weak position and (b) strong position. . . . .	48
9	Preferences depending on the faithfulness relationship between strong and weak geminates. . . . .	49
10	Preferences depending on the output types. . . . .	50
11	Preferences for operation per types of geminates within (a) weak geminates and (b) strong geminates. . . . .	51

# Abstract

## DEGEMINATION IN JAPANESE LOANWORDS FROM ITALIAN

by

Maho Morimoto

In Japanese native phonology, geminate consonants are contrastive (as in [kata] ‘shoulder’ vs. [katta] ‘win-PAST’), but geminates in loanwords can have differing sources and motivations (see Kubozono, Itô, Mester 2009, Kawagoe 2015, and references cited therein): we see gemination of singletons in loanwords from English, in which consonant length is not distinctive ([kæt]<sub>Eng</sub> ‘cat’ → [kjatto]<sub>Jp</sub>), whereas we see geminate-preservation in loanwords from Italian ([espresso]<sub>It</sub> ‘espresso’ → [esupuresso]<sub>Jp</sub>), in which the length of most consonants is contrastive. In loanwords from Italian, however, not all geminates are preserved. This research addresses the cases of degemination, and captures the pattern as stress-based neutralization (Beckman 1998) of consonant length within the framework of Optimality Theory (Prince & Smolensky 1993). Through a database built from dictionaries and a nonce-adaptation survey conducted online, it confirms the preference towards geminates in penultimate position and the ban against geminates in other positions, especially for liquid geminates.

## **Acknowledgments**

My greatest thanks to Armin Mester, Junko Itô, and Grant McGuire for their invaluable input and guidance, as well as their patience. Thanks are also due to the participants of Winter 2015 Research Seminar, Pranav Anand and Martin Krämer for helping me developing this project. I am also grateful to Clara Sherley-Appel, Jeff Adler, and Naoya Watabe for comments, advice, and support, Jed Sam Pizarro-Guevara and Allan Schwade for helping me with the data analysis, as well as Mattia Damaggio for sharing native judgments. All errors are my own.

# 1 Introduction

This paper is concerned with the weight preservation in Japanese loanwords<sup>1</sup> from Italian, with special focus on the patterns of degemination. While both Italian and Japanese have a native contrast in consonant length, as illustrated in (1), geminates in Italian are known to degeminate when adapted into Japanese, under certain circumstances.

- (1) Japanese [kata] ‘shoulder’ vs. [katta] ‘win-PAST’  
Italian [fato] ‘fath’ vs. [fatto] ‘fact’

It has been statistically observed that Italian geminates are more likely to be preserved when they belong to the last three-syllable window in the adapted forms (Tanaka, 2007). This is most obviously illustrated in words including multiple geminates within a word, shown below. Throughout the paper, I use the acute accent mark to indicate both Italian stress-accent and Japanese pitch-accent to ease the comparison. Non-accentedness in Japanese will be indicated by [-].

- (2)
- |    | source      |   | loan      | Italian orthography | gloss             |
|----|-------------|---|-----------|---------------------|-------------------|
| a. | zukkótto    | → | zukótto   | <i>zuccotto</i>     | (a type of cake)  |
| b. | orekkjéttē  | → | orekiéttē | <i>orecchiette</i>  | (a type of pasta) |
| c. | kaffellátte | → | kaferátte | <i>caffè latte</i>  | ‘cafe latte’      |

The primary goal of this paper is to investigate the motivating factors of this positional effect on degemination and to propose a formal analysis within the framework of standard Optimality Theory (OT: Prince and Smolensky 1993), as well as to examine the

---

<sup>1</sup>In this paper, I use “[language name] borrowings” to refer to borrowings *from* the language. For example, “Italian borrowings” refer to words that were borrowed *from* Italian *into* Japanese. The recipient language is always Japanese, unless otherwise indicated. On the other hand, I will use “Japanese loanwords” as a cover term for words that were adapted *to* Japanese from other languages. Throughout the paper, I will be referring to the Tokyo dialect of Japanese unless otherwise indicated.



predictions and ramifications of the hypothesis. I will claim that the asymmetry in the adaptation of geminates can be explained as stress-based positional neutralization, whereby the contrast in consonant length is preferentially preserved in the prominent position in the source form. The flow of the paper is as follows: section 2 lays out the relevant properties of the two languages, and clarifies my assumptions regarding the phonological representation of geminates and the notation that I will be using throughout the paper. Section 3 presents the basic data and my proposal. Section 4 establishes a formal analysis pertaining to my generalization of the data. Section 5 discusses the variation in the adapted forms. Section 6 is a report on an online nonce-word adaptation study, and section 7 concludes.

## 2 Background

### 2.1 Italian and Japanese Phoneme Inventories

The building blocks of Italian and Japanese are strikingly similar. The following charts represent the Italian and Japanese vowel inventories:

(3) Italian and Japanese vowel inventories (Krämer, 2009; Labrune, 2012)

a. Italian		
	–back	+back
+high	i	u
	e	o
	ɛ	ɔ
+low		a

b. Japanese		
	–back	+back
+high	i	u
	e	o
+low		a

The Japanese high back unrounded vowel, [u], will be indicated as [u] for simplicity. Italian mid lax vowels [ɛ, ɔ] are only contrastive in stressed syllables, and will be simplified to [e, o] in this paper to emphasize their correspondence to Japanese [e] and [o].

The consonantal phoneme inventories are also similar between these two languages. In (4), we see that the Japanese consonant inventory is almost the subset of Italian inventory, except for [h] and [ŋ].

(4) Consonant inventories for Italian<sup>2</sup> and Japanese<sup>3</sup>

a. Italian					b. Japanese					
	Labial	Coronal	Palatal	Velar	Labial	Coronal	Palatal	Velar	Uvular	Glottal
Stop	p, b	t, d		k, g	p, b	t, d		k, g		
Affricate		ts, tʃ	tʃ, tʃ			ts	tʃ			
Fricative	f, v	s, z	ʃ, (ʒ)		f, v	s, z	ʃ, ʒ			h
Nasal	m	n	ɲ		m	n			ɴ	
Lateral		l	ʎ							
Rhotic		r				r				
Glide			j	w			j	w		

The following table clarifies my assumptions about how an Italian consonant that does not have a direct Japanese corresponding phoneme is adapted, along with its notation that will be used in this paper:

(5) Italian consonant adaptation into Japanese

Italian		Japanese
ɖ	→	z
ɟ	→	ɟ
l	→	r
ɲ	→	ɲj
ʎ	→	rj

<sup>2</sup>Adapted from Krämer (2009, p.48): parentheses indicate a phoneme only present in loanwords.

<sup>3</sup>Adapted from Labrune (2012, p.60): [ɸ, β] are simplified to [f, v]. I follow Vance (2008) on the uvular (rather than velar) articulation of syllable final nasals.

## 2.2 The Status of Geminate in Italian and Japanese

### 2.2.1 Geminate in Italian

Consonant length is distinctive in both languages. This subsection briefly summarizes the phonological status and relevant phonetic properties of geminates in Italian and in Japanese. Representation of geminates in respective orthographic system will also be discussed.

In Italian, consonant length is distinctive for the most part, while vowel length is predictable.<sup>4</sup>

(6) Minimal pairs from Esposito and Di Benedetto (1999)

- a. [papa] ‘pope’ vs. [pappa] ‘baby food’
- b. [rito] ‘rite’ vs. [ritto] ‘stand up’
- c. [pala] ‘shovel’ vs. [palla] ‘ball’

Some of the consonants occur only as geminates word-internally: all obstruents can occur as long, and [ts, dz, ʃ, ɲ, ʎ] are always long word-internally. There is no long [z], and [r] is realized as tap when short and as trill when long (Krämer, 2009).

The acoustic correlates of geminates in Italian have been examined in works such as Esposito and Di Benedetto (1999) for plosive consonants. According to the experimental results, the duration of the consonant is the primary acoustic correlate, with some contribution of the shortening of the preceding vowel. Reduction of the following vowel is also suspected to be a contributing factor, but they did not find a significant effect regarding this measure.

---

<sup>4</sup>Consonant length is predictable given certain syntactic conditions (*raddoppiamento sintattico*), but they will not be considered in this paper: we consider only word-internal geminates.

### 2.2.2 Geminates in Japanese

In Japanese, consonant length is distinctive, as well as vowel length. Geminates are traditionally called *sokuon*, and are one of the few types of coda allowed in Japanese syllable structure.

(7) Minimal pairs from Kawagoe (to appear)

- a. [ittai] ‘party’ vs. [itai] ‘corpse’
- b. [iffi] ‘one child’ vs. [ifi] ‘volition’
- c. [ippai] ‘one defeat’ vs. [ihai] ‘mortuary tablet’

Geminates are abundant in the Japanese vocabulary, but the distribution varies among the lexical strata. Japanese is known for its lexical organization consisting of the Native (or Yamato), Sino-Japanese, Mimetics, and Foreign (or Western loans) strata (Martin, 1952; McCawley, 1968; Vance, 1997; Shibatani, 1990; Itô and Mester, 1993; Kubozono, 1995; Itô and Mester, 1999a, among others). The Native and Sino-Japanese items allow only for voiceless geminates. Note that as can be seen in (7c), geminated [h] usually surfaces as [pp], except for certain German loans such as *Bach* → [bahha], or *Zürich* → [tʃu:rihhi]. On the other hand, in the Foreign stratum, in which geminates are said to be extremely common, voiced obstruents are allowed (Kubozono et al., 2009; Kawagoe, to appear), while they are subject to devoicing under certain circumstances (Nishimura, 2003, 2006; Sano and Kawahara, 2013). Furthermore, as Tanaka (2007) characterized, Italian loans may include sonorant geminates,<sup>5</sup> which are prohibited in most of the lexical strata (Kawahara, 2005). Meanwhile, long nasals are allowed in Japanese native phonology, and is traditionally referred to by using the notion of *hat-*

---

<sup>5</sup>The articulation of liquid geminates is left for further investigation. Italian rhotic, adapted as tap when single, seems to be adapted as long lateral or a tap following a glottal stop (impression of the author) when geminated, in borrowings such as [farufarre] *farfalle*. If the geminate liquid is realized as a glottal stop, it follows Vance (2008)’s generalization about sonorant geminates.

*suon*, or coda nasal. Geminated nasals in Italian can be either preserved or degeminated through adaptation, just like other sounds:

(8)	source	loan	Italian orthography	gloss
	a. pannakótta	→ pannakótta	<i>pannacotta</i>	(a type of dessert)
	b. ap̃polótti	→ annyorótti	<i>agnolotti</i>	(a type of pasta)

In this paper, I will only consider obstruent geminates, excluding nasal geminates.

Like Italian geminates, the primary acoustic correlate of Japanese geminates is the constriction duration. However, the preceding vowels are longer than when preceding singletons, which differs from the pattern in Italian (Kawahara, to appear). Other secondary cues include the reduction of the vowels following the geminates, increased intensity, larger F0 drop, lower F1, and smaller spectral tilt.

### 2.2.3 Geminates in Orthographic Systems

In both languages, geminates are orthographically quite transparent. The Japanese kana syllabary includes a letter, ‘ツ’ (small version of katakana for ‘ツ’, [tsu]), for the coda portion of a geminate (there is also the hiragana version ‘っ’, but ‘っ’ is usually the ones used for loanwords). In Italian, geminates are usually indicated by reduplicating the letter for the geminated consonant. Lateral approximant [λλ] and alveolar fricative [ʃʃ] are exceptions:

(9)	orthography	pronunciation
	a. tagliatelle	[taλλatelle]
	b. prosciutto	[proʃʃutto]

Japanese speakers are also familiar with representation of geminates in roman alphabets, in either *Hepburn* romanization or *Kunrei* romanization. They differ from

each other in some details such as the representation of long vowels, but are consistent in representing a moraic consonant by reduplicating the letter for the following consonant (Vance, 2008). (10) exemplifies the orthographic representation of corresponding words in each system:

(10)	Italian spelling	Japanese kana spelling	romanization
	a. grappa	グラッパ	gurappa
	b. espresso	エスプレッソ	esuppresso

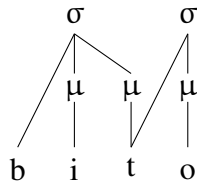
When typing Japanese on a computer keyboard, two of the most popular options are *kana* input mode and *romaji* (or roman alphabets) input mode. In *kana* input mode, each key is assigned one letter from the *kana* syllabary. In order to input ‘っ’, it is necessary to press the shift key and the key for ‘つ’ (hiragana version of ‘ツ’) at the same time. The *romaji* input mode mostly follows the romanization systems, and the combination of ‘k’ and ‘a’ types the *kana* syllabary for [ka]. Today, the *romaji* input mode using a QWERTY keyboard is assumed to have greater share, and is part of the compulsory education starting third grade (Ministry of Education, Culture, Sports, Science, and Technology, 2008).

The next subsection illustrates the phonological representation of geminates that I will be assuming throughout the paper.

### 2.3 Phonological Representation of Geminates

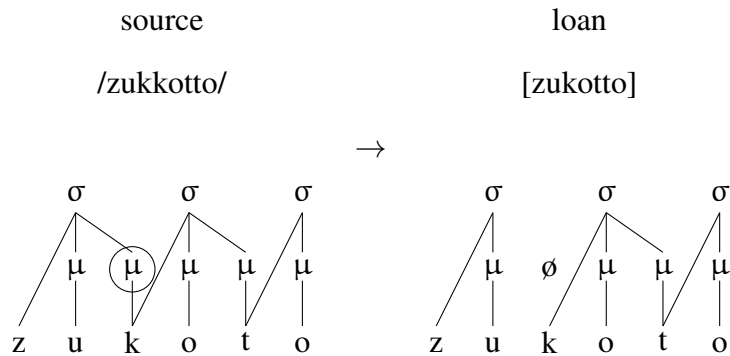
I take up the moraic theory of syllable weight, and assume (11) as the phonological representation of geminated consonants. A geminated consonant is associated with two syllables at the same time, both as an onset and a coda, through a mora linked to the earlier syllable.

- (11) Phonological representation of *bitto*, [bitto] (a type of cheese)



Thus the process of *degemination* or consonant shortening can be seen as the delinking and deletion of the mora associated with a consonant. This is illustrated in the case of *zuccotto* below, where geminate [t] is preserved while geminate [k] undergoes degemination.

- (12) Degemination of [k] in *zuccotto*



In this paper, the first part of geminates will be indicated using capital letters, to make the geminate preservation and simplification visually explicit. In addition, syllable boundaries will be indicated with period [.]. Thus, [zukkotto] will be indicated as [zuK.koT.to], and [zukotto] will be [zu.koT.to].



### 3 The Basic Data and the Proposal

The occurrence of geminates in Italian borrowings is at least partially driven by its existence in the source forms (Tanaka, 2007). In other words, it is the result of the realization or preservation of a mora existing in the input. This can be contrasted with the occurrence of geminates in English borrowings, whereby singletons are geminated in the process of borrowing:

- (13) Geminata in English borrowings
- a. kæt → kjaT.to ‘cat’
  - b. tɒp → toP.pu ‘top’
  - c. pɪk.nɪk → pi.ku.niK.ku ‘picnic’

In (13), a word-final coda consonant preceded by a checked (or lax) vowel is geminated (Ohye, 1967; Lovins, 1975; Kawagoe and Arai, 2002, among others). There have been various views as to the motivation for such gemination, including the preservation of the coda status of the consonant (Katayama, 1998, among others) and the effect of Japanese prosodic constraints (Kubozono, Itô, and Mester, 2009). Kubozono, Itô, and Mester (2009) draws attention to the positional effect, whereby gemination tends to occur word-finally, as exemplified in (13c). The gemination process in English borrowings is sensitive to segmental conditions that are quite close to that of the native ones. Voiced segments are rarely geminated, while there are exceptions such as *dog* → [doG.gu], *bag* → [baG.gu]; [s, f] do not undergo gemination while [ʃ, x] do (Kubozono, Itô, and Mester, 2009); sonorant obstruents do not undergo gemination.

On the other hand, Tanaka (2007) and Tanaka and Kubozono (2008) demonstrate that the appearance of geminates in Italian borrowings is due to preservation, or in OT terms, the effect of highly ranked faithfulness constraints. The adaptation of Italian

geminate is illustrated below:

(14)	source	loan	Italian orthography	gloss
a.	bíT.to	→ bíT.to	<i>bitto</i>	(a type of cheese)
b.	es.préS.so	→ e.su.pu.réS.so	<i>espresso</i>	(a type of coffee)
c.	far.fáL.le	→ fa.ru.fáR.re	<i>farfalle</i>	(a type of pasta)
d.	gráP.pa	→ gu.ráP.pa	<i>grappa</i>	(a type of drink)
e.	kar.páT.tjo	→ ka.ru.páT.tjo	<i>carpaccio</i>	(a type of appetizer)
f.	ji.róK.ko	→ ji.róK.ko	<i>scirocco</i>	‘south-east wind’
g.	pan.tjéT.ta	→ pan.tjéT.ta	<i>pancetta</i>	(a type of bacon)
h.	for.tíS.si.mo	→ fo.ru.tíS.ji.mo	<i>fortissimo</i>	(musical term)
i.	vja.reD.ɕo	→ bi.a.reD.ɕo	<i>Viareggio</i>	(place name)
j.	o.pe.ra.buF.fa	→ o.pe.ra.buF.fa	<i>opera buffa</i>	(a form of opera)

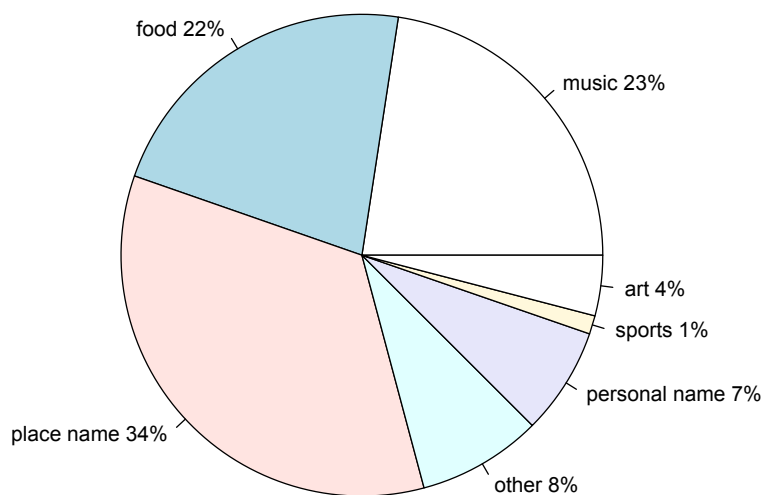
Italian borrowings seldom undergo gemination, as the Italian syllable structure does not allow word-final coda obstruent in the first place. There are exceptions such as *rucola* → [ruK.ko.ra], *bufala* → [buF.fa.ra], or *amatriciana* → [a.ma.to.riT.tʃa.na], in which Italian singletons undergo gemination, but they are very few. The claim that the appearance of geminates in Italian is based on their presence in the source form is further supported by the laxed segmental condition: [s, f] and sonorant geminates can be preserved (14b, h, j), as well as voiced obstruents (14i), while not as common as voiceless obstruents.

The puzzle here is that not all geminates in Italian are preserved in their loan forms. Assuming that preservation of Italian geminates in their Japanized forms is motivated by faithfulness constraints, we expect all geminates to be preserved. However, this expectation is not borne out. In the next subsection, I examine the positional effect on the preservation and simplification of geminates.

### 3.1 The Data

Italian borrowings in Japanese are not as numerous as English borrowings, and Irwin (2011) points out the possibility that they were mediated by other languages Japanese has been primarily borrowing from, but they have been contributing in the culinary and musical domains in the modern period. For example, Irwin (2011) mentions *opera* → [opera] in 1881, *salami* → [sarami] in 1901, *spaghetti* → [supagetti] in 1931, and *pasta* → [pasuta] in 1975. Figure 1 is a chart indicating the domains of loanwords from Italian in my database, showing the prevalence of terms related to food and music besides place names.

Figure 1: Genres of Italian borrowings in Japanese



The data presented in this paper, unless otherwise indicated, are taken from the database that I built out of seven dictionaries of Japanese (*Kōjien* (Shinmura, 1998);

*Japanese pronunciation accent dictionary* (NHK Hōsō Bunka Kenkyūjo, 1998); *Concise katakana go jiten* (Sanseidō Henshūjo, 2010); *Super Daijirin* (Sanseidō Henshūjo, 2015); *Shinmeikai kokugo jiten* (Yamada et al., 2011); *Concise foreign place name dictionary* (Tanioka, 1998); *Daily concise Japanese dictionary* (Sanseidō Henshūjo and Satake, 2010)) and an Italian-Italian dictionary, *Zingarelli* (Zanichelli Editore Spa, 2013)<sup>6</sup>. I extracted 1209 words borrowed from Italian (including personal names and place names). Two different adaptation forms for a single source form were separately counted (for example, the Italian form *adagio* contributes two data points to the database, [ada:ç̥o] and [adaç̥io]). The database is meant to be a recreation of Tanaka (2007)'s corpus study, in which he acknowledged the positional effect on geminate preservation in Italian borrowings.

Within the 1209 words, there were 5059 occurrences of Italian consonants (a long consonant was counted as one occurrence of the consonant), whereby 526 were geminates (the number of geminates including nasal geminates was 569). On the other hand, the occurrence of *sokuon*, or obstruent geminates in Japanese, was 305. The overall rate of geminate preservation was 57% (5 of the geminates in the adapted forms were formed through gemination, not originating from Italian geminates). The table in (15) summarizes the rate of geminate preservation for each segment, following Tanaka (2007)'s style, in order to show the segmental effect on the preservation of geminates.

In (15), it is clear that voiceless obstruent geminates are predominantly preserved. Voiced obstruent geminates are not as numerous in the source forms (about a tenth of the voiceless obstruent geminates in Italian), but some of them are preserved, especially [dç̥]. Liquid geminates are much popular in the source forms, and about a third of them are preserved in the adapted forms.

---

<sup>6</sup>Glossaries are making use of Wikipedia and Wiktionary entries.

(15) Geminate preservation rate in my database<sup>7</sup>

Voiceless Obs.			Voiced Obs.			Liquid		
pp	81%	(26/32)	bb	14%	(1/7)	ll	30%	(32/108)
tt	85%	(81/95)	dd	100%	(1/1)	rr	23%	(7/30)
kk	68%	(30/44)	gg	-		ʎʎ	17%	(4/23)
tts	53%	(20/38)	ddʒ	50%	(1/2)			
ttʃ	77%	(27/35)	ddʒ	70%	(16/23)			
ss	76%	(47/62)	vv	-				
ʃʃ	40%	(4/10)						
ff	19%	(3/16)						
72% (238/332)			58% (19/33)			27% (43/161)		

In addition to the segmental conditions, positional effect on the preservation of Italian geminates is also observed. (16) summarizes the distribution of geminates in the adapted forms by syllable position, in my database. The syllable counting is done on the nativized form rather than the Italian form. There is no geminates in the ultimate syllable, as neither language allows for word-final geminates.

## (16) Distribution of geminates by position

Position	further left	5th	4th	antepenultima	penultima	total
Geminate occurrence	10%	7%	14%	24%	45%	100%

About half of the preserved geminates are located in the penultimate syllable. Tanaka (2007), in his dictionary-based research, looked at the preservation rate for each syllable position (he also looked at the word-length effect, but it does not seem to be as robust as the positional effect, and will not be discussed in this paper).

<sup>7</sup>The preservation rates of Italian [s] and [ʃ] are tricky, as some [s] are adapted as [ʃ] depending on the following vowel. In the adapted forms, the occurrence of [s] was 31 times and [ʃ] was 20 times.

## (17) Preservation rate of geminates by position

Position	further left	4th	antepenultima	penultima	total
Preservation rate	29%	38%	60%	73%	60%
Counts	7/24	22/58	62/104	123/169	214/355

In addition to the high preservation rate in the penultimate syllable, we also notice a huge drop between the rates in the 3rd and 4th syllables from the right edge. There are two main generalizations he draws from these results:

- (18) a. Geminates in the source forms are more likely to be preserved in the penultimate syllable in the nativized forms.  
 b. Geminates are more likely to be degeminated when they are outside of the last three-syllable window of the nativized form.

The following data points further exemplify this asymmetry in the adaptation of geminates dependent on position:<sup>8</sup>

(19)	source	loan	Italian orthography	gloss
a.	zuK.kóT.to	→ zu.kóT.to	<i>zuccotto</i>	(a type of cake)
b.	aL.le.gréT.to	→ a.re.gu.réT.to	<i>allegretto</i>	(musical term)
c.	teR.ra.kó T.ta	→ te.ra.kóT.ta	<i>terracotta</i>	(a type of earthenware)
d.	taL.lja.téL.le	→ ta.ri.a.téR.re	<i>tagliatelle</i>	(a type of pasta)
e.	o.reK.kjéT.te	→ o.re.ki.éT.te	<i>orecchiette</i>	(a type of pasta)
f.	aR.raB.bjá:ta	→ a.ra.bi.á:ta	<i>arrabbiata</i>	(culinary term)

In (19), we see that geminates in the penultimate syllable are preserved, while we see

<sup>8</sup>In Zingarelli, *zuccotto* is listed as [tsukkóto]. [ts~z] is dependent on dialects, and I assume [z] throughout, based on the northern dialect which usually pronounces word-initial [z] as [dʒ] (Krämer, 2009, p. 9).

degemination elsewhere. These items are especially of theoretical interest, as they include both instances of geminate preservation and degemination within a word. Because of forms such as (19a, e), and not \*[zúK.ko.to] or \*[o.reK.kí.e.te], we are compelled to look further than just ‘final three-syllable window’ as privileged position for geminate preservation.

### 3.2 The Proposal

In what follows, I claim that the preservation of geminates in Italian loanwords is dependent on the prominence in the source language. As it turns out, the last three-syllable window, in which geminates are preferentially preserved, is exactly where the Italian stress falls onto,<sup>9</sup> and it is also where the Japanese pitch accent is assigned to in loanwords,<sup>10</sup> while not always. Going back to the data points in (19) and the tendencies clarified in Tanaka (2007), they seem to suggest that geminates in these prominent positions tend to be protected. What needs to be clarified, then, is whether it is the prominence in Italian or the prominence in Japanese that decides whether a geminate should be protected or not. The fact that the prominent positions in the source and recipient languages converge most of the time makes it difficult for us to know which is responsible for the privilege. In his conclusion, Tanaka (2007) mentions that it is meaningful to keep the geminate in a position where accent is most-likely assigned in Japanese loanwords. While I am not to deny the force of the prosody of the recipient language or the possibility that the prosody of both languages works together to yield the positional effect, I pursue a scenario in which the prominence of the source lan-

---

<sup>9</sup>Italian stress assignment on nouns are thought to be lexically specified, but most of them falls on the penultimate syllable (Krämer, 2009, chpt. 6)

<sup>10</sup>Loanwords tend to be accented on the syllable containing the antepenultimate mora (Martin, 1952; McCawley, 1968; Kubozono, 2006; Itô and Mester, 2014, among others) which often ends up on the penultimate syllable (especially in Italian borrowings, as Italian stress attracts weight).

guage is responsible for the privileged status of geminates in the penultima. Below, I would like to present data points that are indicative of the positional effect being dependent on the head status of the syllables in the source form. The few cases where the source and loan prominence diverges allow us to arbitrate between the two (not necessarily mutually exclusive) possibilities. First, a geminate in a non-stressed syllable is not necessarily preserved when it is adapted in an accented syllable in Japanese. This is shown in (20), where accented syllables in the loan forms are boxed.

(20)	source	loan	Italian	gloss
a.	suP.plí	→ <span style="border: 1px solid black; padding: 0 2px;">sú</span> .pu.ri	<i>suppli</i>	(a type of rice croquette)
b.	piK.ká:ta	→ <span style="border: 1px solid black; padding: 0 2px;">pí</span> .ka.ta	<i>piccata</i>	(a dish of escalope)
c.	maT.téo	→ <span style="border: 1px solid black; padding: 0 2px;">má</span> .te.o	<i>Matteo</i>	(personal name)

What we see above is the fact that the accented status of a syllable in the nativized form does not necessarily protect the moraicity of a consonant. The boxed syllables in the loan forms are the locations we expect geminate preservation, were it dependent on the assignment of Japanese pitch accent, but the geminates are nonetheless simplified. Furthermore, (21) shows cases in which the moraicity of a consonant dominated by a stressed syllable in the source form is preserved, when the prominence of the syllable has been dislocated through the loan adaptation process. Below, syllables in the loan forms that correspond to the stressed syllables in the source forms are boxed.



(21)	source	loan	Italian	gloss
a.	fal.séT.to	→ fá.ru.seT.to	<i>falsetto</i>	(musical term)
b.	kwin.téT.to	→ ku.ín.teT.to	<i>quintetto</i>	(musical term)
c.	dju.éT.to	→ djú.eT.to	<i>duetto</i>	‘duet’
d.	ri.sóT.to	→ rí.zoT.to	<i>risotto</i>	(a dish of rice)
e.	píK.ko.lo	→ piK.ko.ro-	<i>piccolo</i>	(musical instrument)
f.	fa.góT.to	→ fá.goT.to	<i>fagotto</i>	(musical instrument)

What we see in (21) is the fact that the preservation of geminates is not necessarily dependent on the accented status of the syllable it belongs to in the loan form. Rather, it is the membership to a stressed syllable that enhances the preservation of a geminate. One may still argue that the preservation is simply dependent on the membership to the penultimate syllable in the loan form. However, this is not compatible with (21e), where the geminate in the stressed syllable is kept when it does not belong to the penultimate syllable or in an accented syllable in Japanese.<sup>11</sup>

Given these data points, I propose that the positional effect on degemination is dependent on the head status of the syllable to which the mora of a geminate consonant belongs in Italian. Before a serious and formal pursuit of this proposal, however, certain issues need to be addressed, along with my assumptions as to the transient and gradient nature of loanwords.

First, it should be acknowledged that the data points in (20) and (21) are not the most crucial ones – considering the nature of loanwords discussed below, we cannot conclude that it is the Italian prominence and not the Japanese prominence that is in

<sup>11</sup>These crucial cases are, however, not easy to find. The possibility that they are lexical exceptions cannot be dismissed at this point. Also thanks to Naoya Watabe for pointing out that some of the adaptation forms in (21) can be accented differently: the loan form for *duetto*, can be unaccented in its loan form [dju.eT.to-]; the loan form for *risotto* can be accented in the penultimate syllable as in [ri.zóT.to].

play, until we come across a source-adaptation pair that contains two or more geminates in the Italian form, of which only one is degeminated, and the syllable containing the preserved geminate is stressed, but not accented in its adapted form. Schematically, such a pair would look like:

(22) /c'vC.cvC.cv/ → [cvC.cv.cv-]

Meanwhile, this or a similar pattern may not be viable in Italian nouns, as heavy penultimate attracts stress (Krämer, 2009, chp. 6). In that case, it would be interesting to test it in a perceptual study with either nonce-words, or with other categories such as verbs or adjectives.

The reason why we could not overestimate the impact of (20) and (21) on our generalization pertains to the aforementioned structure of the phonological lexicon in Japanese, or to the fact that loanwords from Italian are not monolithic in the Japanese lexicon. Assuming a core-periphery structure rather than an aggregate of sublexicons (Itô and Mester, 1995a,b, 1999a), it is plausible that the items in (20) and (21) belong to a strata that either preserves or degeminates geminates across the board, the former being closer to the periphery and the latter being closer to the core. The importance of the data in (19) was due to the fact that these items do not belong to either of them, but instead in a strata in which the nativization is not complete. Because it is an intermediate stage in the process of nativization, it allows us to see constraint interactions that would not be apparent otherwise.

In the face of this lack of definitive piece of data, however, referring to the prosody of a foreign input can be supported by other phenomena. I would like to abstract away from committing to a precise mechanism through which foreign words are borrowed into a language (and as Kawagoe and Arai (2002) note, there are many possible sce-

narios as to the process in which loanwords are adapted). However, it seems clear that the prominence in the source language may have impacts on the adaptation forms. According to Kubozono (2006), prominence of the source language can have some effect on the accentedness and the location of the accent in Japanese loanwords. Then, it should not be that weird that Italian prominence can affect the adaptation process into Japanese.

A different kind of support can be made from a phonetic standpoint. An acoustic investigation of Italian geminates revealed that geminates are typically shorter in unstressed conditions (Payne, 2005).<sup>12</sup> While we lack historical or linguistic evidence for the borrowing process to take place via auditory input, there may be perceptual reasons for preserving geminates that are more salient in stressed syllables in Italian (for the role of perceptual saliency in loanword adaptation, see Yip 2002).

In the next section, I further suggest that this can be viewed as positional neutralization based on Italian stress position, and that the effect can be formalized using the positional faithfulness schema, assuming an Output-output correspondence relationship between the source and adapted forms, within standard Optimality Theory (OT: Prince and Smolensky 1993).

---

<sup>12</sup>There is a complication on the differences in the duration between post-stress and pre-stress geminates, which is dependent on the sentence position (nuclear vs. non-nuclear). This cannot be discussed further in this paper.

## 4 Theoretical Assumptions and the Basic Analysis

In this section, I cover the two theoretical assumptions that I will be making in explaining the positional effect on degemination in Japanese loanwords from Italian.

### 4.1 Output-output Faithfulness

Taking up the assumption that the positional effect on degemination is dependent on the head status of the syllable it belongs to in Italian, the input has to be fully prosodically specified. Moreover, as we saw in (21), the input to the loan process has to be prosodified according to the Italian phonology. This is given in (23):

(23)	Italian output (input)	Japanese output (output)
	/zuK.kóT.to/ <sub>It</sub>	→ [zu.kóT.to] <sub>Jp</sub>
	/o.reK.kjéT.te/ <sub>It</sub>	→ [o.re.ki.éT.te] <sub>Jp</sub>
	/maK.kjá:to/ <sub>It</sub>	→ [ma.ki.á:to] <sub>Jp</sub>

In my analysis, I adopt the assumption that the input to the Japanese loan process is the output of Italian phonology. The Italian output and Japanese output will be referred to as if they were in the usual Input-output correspondence in OT henceforth (more on Output-output correspondence in McCarthy and Prince 1995). This assumption is also supported by earlier works in loanword phonology within the framework of OT, many of which assuming this kind of Output-output correspondence: it is not an uncommon assumption as exemplified in Greek stress pattern in Revithiadou (1999), Thai vowel length in Kenstowicz and Suchato (2006), SB-FAITH constraints in Smith (2006), and so on. It should be noted that this assumption does not necessarily commit us to the idea that all Japanese speakers store loanwords from Italian in their foreign output forms, or to the idea that all Italian loanwords have been borrowed by a bilingual speaker of

Italian and Japanese.

## 4.2 Positional Faithfulness

The positional effect on degemination can be captured as neutralization of consonant length outside a prominent position, and the asymmetry can be obtained in OT by introducing the general schema following Beckman (1999):

(24) Positional neutralization ranking schema

IDENT-*Position*[F] » M » IDENT[F]

In this schema, the baseline is the neutralization rendered by the ranking M » IDENT[F]. The ranking IDENT-*Position*[F] » M protects a property in the specific position from neutralizing. Beckman (1999, ch. 3) discusses stress-based positional neutralization, in which the inventory of contrastive elements in unstressed syllable is a subset of the one in stressed syllable. The case of degemination in Italian borrowings can be viewed as an instance of stress-based neutralization of consonant length, in which the contrast surfaces only in the stressed syllable. The twist here is that the accented status of the syllable does not necessarily surface, as the accent is presumably re-assigned according to the Japanese prosody.<sup>13</sup> Also, while Beckman's (1999) constraint looks at an output segment in a stressed syllable to seek its correspondent in the input, our positional faithfulness constraint, as formulated below in (25), takes an output segment to check its prominent status and its specification in the input. Question remains as to

---

<sup>13</sup>Thanks to Nick Kalivoda for pointing out that this is a case of opacity, and that the reference to the prominent status of the syllable occurs in a slightly different way from Beckman. Meanwhile, assuming that the positional effect is dependent on the Japanese prominence, things are going to look much more like her positional faithfulness, calling for faithfulness constraints that refer to the surface prosodic properties. However, then, it may be hard to provide an account in parallel OT, as the loan process has to decide on the accentuation and preservation of geminates which is dependent on the accentuation at the same time.

the legitimacy of such a constraint, but to get started, I propose the following family of constraints that requires a segment to be consistent on its moraic specification in the input and the output:

(25) IDENT-' $\sigma[\mu]$  : let  $\beta$  be an input segment in a stressed syllable, and  $\alpha$  its output correspondent. The moraic value of  $\alpha$  must be identical to that of  $\beta$ .

“An input segment in a stressed syllable and its output correspondent of that segment must have identical moraic specifications (weight and association).”

(26) IDENT[ $\mu$ ]: let  $\beta$  be an input segment and  $\alpha$  its output correspondent. The moraic value of  $\alpha$  must be identical to that of  $\beta$ .

“An input segment and the output correspondent of that segment must have identical moraic specifications.”

(27) Subhierarchy of the faithfulness constraints

IDENT-' $\sigma[\mu]$  » IDENT[ $\mu$ ]

The general schema of these IDENT constraints is adapted from McCarthy and Prince (1995), assuming a theory in which a mora is treated as a weight property, much like other features (they do not directly address the possibility of constraining moraic association using this constraint). Thus the constraint inherits most of Beckman (1999) 's design. The definition of this constraint also includes some flavor of IDENT-R[ $\mu$ ] in Itô and Mester (1999a) and Itô and Mester (2003), so that it makes monomoraicity/bimoraicity distinctions. In that sense, it is equivalent to IDENT-' $\sigma$ LENGTH, reminiscent of IDENT-LENGTH( $\sigma_1$ ) in Brennan (2006) suggested for Latin.

An alternative is to appeal to moraic faithfulness constraints (see more in Morén 2001, sec.2.2.3) such as MAXLINK-' $\sigma[\mu]$  and DEPLINK-' $\sigma[\mu]$ , treating a mora as a prosodic entity rather than a property. They yield the same effect as IDENT-' $\sigma[\mu]$ , in

conjunction. A similar effect can be obtained by  $\text{MAX-}'\sigma[\mu] \gg \text{MAX}[\mu]$ , but  $\text{MAX}$  alone cannot distinguish between candidates with geminate preservation and compensatory lengthening: in section 5, we will need to distinguish between faithfulness to the moraic specification and faithfulness to quantity, or  $\text{IDENT-}'\sigma[\mu]$  and  $\text{MAX-}'\sigma[\mu]$ .

The subhierarchy in (27), when intervened by a markedness constraint that militates against geminates, renders the positional effect on degemination: all geminates in a word will be degeminated except the one whose mora belongs to the stressed syllable in the input. I define the markedness constraint as in (28), to give the subhierarchy in (29):

(28)  $\text{NOGEM}$ : assign a violation for each consonant that is a geminate.

(29)  $\text{IDENT-}'\sigma[\mu] \gg \text{NOGEM} \gg \text{IDENT}[\mu]$

This constraint implies that geminates are universally marked than singletons. It would be appropriate to assume more markedness for geminates than for singletons, given that not all languages are endowed with geminate consonants. I follow Kubozono, Itô, and Mester (2009) in assuming  $\text{NOGEM}$ . Possible alternatives include  $*\mu/C$  which prohibits a consonant to be moraic (thus it penalizes all coda consonants, including nasals<sup>14</sup>) and  $\text{CRISPEGE}$  (Itô and Mester, 1999b), adopted in D’Imperio and Rosenthal (1999) to penalize geminates in that a consonant for being dominated by two syllable nodes.

### 4.3 Deriving the positional effect in *zuccotto*

In this subsection, I show how the constraints and ranking in (29) can derive the positional effect on degemination, in a very basic case of *zuccotto*,  $/\text{zuK.kóT.to}/ \rightarrow$

<sup>14</sup>I thank Naoya Watabe for pointing out that  $*\mu/C$  does not penalize word-initial geminates, or any other geminates that do not involve moraic consonants as assumed in section 2.3.

[zu.kóT.to].

(30)

	/zuK.kóT.to/	IDENT-' $\sigma$ [ $\mu$ ]	NOGEM	IDENT[ $\mu$ ]
a.	zu.koT.to		*	*
b.	zuK.koT.to		**!	
c.	zu.ko.to	*!		**
d.	zuK.ko.to	*!	*	*

Candidates with degemination in the prominent position (30c,d) are immediately ruled out by IDENT-' $\sigma$ [ $\mu$ ]. Candidates preserving the geminate in the prominent position (30a,b) survive IDENT-' $\sigma$ [ $\mu$ ], but the general ban on geminates eliminates candidate (30b). Thus the winner is the positionally faithful candidate (30a).

It is worth noting here that this configuration makes a strong prediction about (30d): candidate (30d) is harmonically bounded, and cannot be chosen under any of the six possible permutations of these three constraints. I call this candidate an anti-degemination candidate, as the geminate in weak position (or unstressed syllable in Italian) resists the pressure from NOGEM that a geminate in strong position (or stressed syllable in Italian) surrenders.

The same configuration accounts for degemination in other tokens<sup>15</sup> (see (19) for examples), often interacting with other constraints for additional repair strategies, such as vowel epenthesis (/aL.le.gréT.to/ → [a.re.gu.réT.to]) or glide vocalization (/o.reK.kjéT.te/ → [o.re.ki.éT.te]).

<sup>15</sup>One good thing about this constraint is that it works with English loanwords as well: *pet* /pét/<sub>Eng</sub> → [pét.to]<sub>Jp</sub>. However, Kubozono, Itô, and Mester (2009) 's analysis still holds in that Japanese prosodic constraints can override this faithfulness constraint as in *doctor* /dók.tə/<sub>Eng</sub> → [dók.ku.ta:]<sub>Jp</sub>.



## 5 Variation

The constraints and ranking established in the previous section makes the prediction that any geminate consonants in Italian borrowings, when not belonging to a stressed syllable in the input, would be categorically degeminated. The data do not quite conform to this generalization, in at least three ways:

- (31)
- a. There are many cases in which degemination does not take place, regardless of its position.
  - b. Simple degemination is not the only way to repair for geminates: degemination accompanied with compensatory lengthening is also an option.
  - c. The categorical nature of the positional degemination does not capture the hierarchy among different types of geminates.

This section discusses these issues.

### 5.1 Free Variation

To understand the first point raised above, it is useful to take a look at the way in which an Italian form can often correspond to multiple loan forms in usage. This is illustrated in *zuccotto* having multiple attested forms: losing candidates in the tableau (30) for *zuccotto* are in free variation, as can be seen in (32) using Google search hits as an informal indicator of their frequency of usage (as of February 20, 2015).

(32)

Adapted Forms	Google Hits
a. zu.koT.to	203,000
b. zuK.koT.to	1,870
c. zu.ko.to	349
d. zuK.ko.to	109

The overwhelming popularity of the positionally faithful form (32a) is in line with our prediction in the previous section. Moreover, the violation profile in general seems to roughly align with the Google hit results. This is shown in (33), a replication of tableau (30) with the addition of Google hit results in the rightmost column.

(33)

	/zuK.kóT.to/	IDENT-' $\sigma$ [ $\mu$ ]	NOGEM	IDENT[ $\mu$ ]	Google Hits
☞ a.	zu.koT.to		*	*	203,000
b.	zuK.koT.to		**!		1,870
c.	zu.ko.to	*!		**	349
d.	zuK.ko.to	*!	*	*	109

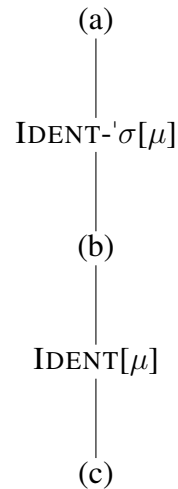
In this tableau, a candidate that violates a higher-ranked constraint are relatively less frequently used. Especially, the harmonically bounded (33d) is also the least frequent output.<sup>16</sup>

This variability in usage looks like an intersection of theories of variation (Anttila, 1997; Boersma and Hayes, 2001; Legendre et al., 1990, to name a few), degree of nativization (Kiparsky, 1968), and the structure of phonological lexicon (Itô and Mester, 1995a,b, 1999a).

<sup>16</sup>The fact that it is actually in usage is out of the scope of my account, but I suspect there are other constraints in play, perhaps a prosodic requirement such as PARSE $\sigma$  in conjunction with other constraints.

The permutation of these constraints gives us different outputs. I assume that IDENT-' $\sigma[\mu]$  always dominates IDENT[ $\mu$ ], and show the three places where NOGEM can be ranked relative to these two faithfulness constraints:

(34) Three possibilities regarding the ranking of NOGEM



Only when we have NOGEM in (b), we get the positional degemination. When we have NOGEM in (c), all geminates would surface without degemination; when we have NOGEM in (a), we see no geminates in the adapted form, and it will be the adaptation most compatible with the Native stratum.

Meanwhile, the focus of the rest of the paper will be on the characteristics of positional degemination, abstracting away from a world in which all Italian geminates are kept (i.e. maximally faithful, very low degree of nativization) or all Italian geminates are degeminated (i.e. maximally satisfying the markedness constraint, high degree of nativization).

## 5.2 Compensatory Lengthening

As mentioned in the second point above, degemination is sometimes accompanied with a lengthening of the preceding vowel, which can be seen as compensatory lengthening. In this subsection, compensatory lengthening in the adaptation of Italian geminates will be illustrated using liquid geminates, as they readily and quite systematically have the vowel-lengthening option.<sup>17</sup>

While sonorant geminates (except for nasals) are generally prohibited in Japanese native phonology, Italian borrowings may preserve sonorant geminates:

(35)	source		loan		Italian	gloss
	a.	ti.ɕéL.le	→	ti.ɕéR.re	<i>tigelle</i>	(a type of bread)
	b.	taL.lja.téL.le	→	ta.ri.a.téR.re	<i>tagliatelle</i>	(a type of pasta)
	c.	sfoL.lja.téL.la	→	su.fo.ri.a.téR.ra	<i>sfogliatella</i>	(a type of pastry)

However, when Italian liquid geminates are degeminated, they have the option of lengthening the preceding vowel, as can be seen below (the greater-than and shorter-than signs indicate the usage frequency according to Google hits):

(36)	source		long vowel		geminate	Italian	gloss
	a.	ta.ráL.li	→	ta.rá:ri	>	ta.ráR.ri	<i>taralli</i> (a type of snack)
	b.	nu.téL.la	→	nu.té:ra	<	nu.téR.ra	<i>nutella</i> ‘nutella’
	c.	ti.ɕeL.le	→	ti.ɕé:re	<	ti.ɕéR.re	<i>tigelle</i> (a type of bread)

The lengthening of the vowel does not co-occur with geminate preservation (it would

<sup>17</sup>Compensatory lengthening is not limited to liquids. Here, we limit our scope to only liquid geminates, but forms such as *arpeggio* and *prosciutto* are also most frequently realized with their vowel-lengthening form, and even [zuK.ko:to] for *zuccotto* is attested. Which segments are more prone to this process is to be confirmed in further investigation, but I expect that the higher the sonority, the more they tend to undergo compensatory lengthening (see Kawahara 2005 for more on sonority and geminates).



(40) NOGEM[R], MAX-' $\sigma$ [ $\mu$ ] » IDENT-' $\sigma$ [ $\mu$ ]

(41) Degemination and vowel lengthening in *taralli* with Google hits<sup>19</sup>

		NOGEM[R]	MAX-' $\sigma$ [ $\mu$ ]	IDENT-' $\sigma$ [ $\mu$ ]	NOGEM	IDENT[ $\mu$ ]	Google Hits
	/ta.ráL.li/						
a.	ta.ra:ri			**		**	32,100
b.	ta.ra.ri		*!	*		*	25,400
c.	ta.raR.ri	*!			*		9,860

In this tableau, the positionally faithful candidate (41c) is not selected as it violates NOGEM[R]. The  $\mu$ -deleting candidate (41b) is not selected as it violates MAX-' $\sigma$ [ $\mu$ ]. Thus (41a), while violating IDENT-' $\sigma$ [ $\mu$ ] twice (penalized twice, once for changing the moraic specification of [r] by delinking, and once for changing the moraic specification of [a] by reassociating), is selected as the optimal output.

The next subsection looks at an item which contains two liquid geminates, of which one is degeminated and the other undergoes compensatory lengthening.

### 5.3 Implicational hierarchy of faithfulness

In this subsection, I try to capture two types of implicational relationships: the hierarchy that exists among geminates in different positions within a word, and the hierarchy that exists among the losing candidates of such a word. This connects to the third point raised at the beginning of this section. The case of *tagliatelle*, which includes two liquid geminates in the input, in strong and weak position, is of special interest as it allows us to see more in these implicational relationships.

<sup>19</sup>[tarari] is a real mimetic word in Japanese, meaning the dripping of some liquid. The Google hit result here is the one for “tarari AND itaria (Italy).” I also omitted some proper nouns including the string [tarari].

(42) Adaptation of *tagliatelle*, replicated from (35)

taL.lja.téL.le → ta.ri.a.téR.re *tagliatelle* (a type of pasta)

Here, a complication is that [j] in the onset of [lja] vocalizes to yield [ri.a]. However, this process (palatal vocalization) is independent of degemination, and will be disregarded in my analysis that follows. Thus I consider nine logical possibilities for geminate adaptation for *tagliatelle*,<sup>20</sup> the three options for each geminate being geminate preservation, degemination, and compensatory lengthening. *weak* and *strong* refer to the positions of geminates, *strong* being the prominent syllable in the source form, and *weak* being everywhere else.

(43) Degemination patterns and frequency (Google hits as of March 11)

Adapted Forms	Google Hits	weak gem	strong gem
a. ta.rja.teR.re	378,148	degem	pres
b. ta.rja.te:re	2,966	degem	comp
c. ta.rja.te.re	2,749	degem	degem
d. taR.rja.teR.re	2,450	pres	pres
e. taR.rja.te:re	44	pres	comp
f. taR.rja.te.re	3	pres	degem
g. ta:rja.teR.re	3	comp	pres
h. ta:rja.te:re	0	comp	comp
i. ta:rja.te.re	0	comp	degem

<sup>20</sup>The actual most frequently attested output form is [ta.ri.a.teR.re], with degemination and palatal vocalization ([j] → [i]) in the weak position. The logical possibilities taking palatal vocalization into account amounts to 3x2x3=18. As palatal vocalization is independent of degemination, I incorporated their Google hits to the candidates seen in (i).

While this is no more than a report based on an informal index on just one data point, it gives us an interesting picture. The foremost thing to note is that in weak position, geminates tend to degeminate. And in strong position, geminates prefer being preserved, otherwise undergoing compensatory lengthening, otherwise degeminating. This seems to point to the following subhierarchy (note that the ranking differs from that of (41), as we are dealing with a case in which compensatory lengthening is not the optimal output):

$$(44) \quad \text{MAX-}'\sigma[\mu] \gg \text{IDENT-}'\sigma[\mu] \gg \text{NOGEM}[\text{R}]$$

The divide between (43a-d) and (43e-i) seems to be indicative of a tendency, such that there cannot be more faithfulness in the weak position than in the strong position. I should keep this hypothesis in a reserved note for the moment, though, because (43g,h) do not have a good excuse for not being popular. But at the bottom line, we do not see as much (43e,f,i), where weak geminates maintain more faithfulness than strong geminates. Schematically, for an input of the form /...cvR.rv...c $\acute{v}$ R.rv/, we do not expect outputs such as [...cvR.rv...cv.rv], [...cvR.rv...cv $\acute{v}$ .rv], or [...cv $\acute{v}$ .rv...cv.rv].

In the following tableau, I combine the configuration in (29) and (44) to see what they predict, with the addition of MAX[ $\mu$ ], assuming the relationship MAX-' $\sigma[\mu]$   $\gg$  MAX[ $\mu$ ] to hold universally. I also assume NOGEM[R]  $\gg$  NOGEM given the relevance of the sonority hierarchy to geminacy (Kawahara, 2005). They correctly predict the victory of a positionally faithful candidate [ta.rja.teR.re].



(45) Positional effect in [ta.rja.teR.re] for *tagliatelle*

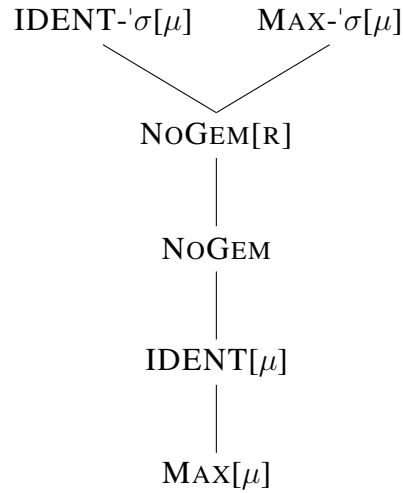
	/taL.lja.téL.le/	MAX- $\sigma$ [\mu]	IDENT- $\sigma$ [\mu]	NOGEM[R]	NOGEM	IDENT[\mu]	MAX[\mu]	Google Hits
a.	ta.rja.teR.re			*	*	*	*	378,148
b.	ta.rja.teɹ.re		*!*			***	*	2,966
c.	ta.rja.te.re	*!	*			**	**	2,749
d.	taR.rja.teR.re			**!	**			2,450
e.	taR.rja.teɹ.re		*!*	*	*	**		44
f.	taR.rja.te.re	*!	*	*	*	*	*	3
g.	taɹ.rja.teR.re			*	*	**!		3
h.	taɹ.rja.teɹ.re		*!*			*****		0
i.	taɹ.rja.te.re	*!	*			***	*	0

In the tableau, we see that (45e,f,i), the anti-degeminating candidates (degeminating the strong geminates while preserving the weak ones) are harmonically bounded, which seems to be what we want. However, unlike (33), there are mismatches between the violation profile and usage. First, preservation of the strong geminate is preferred in the tableau, as can be seen for (45a,d,g), but the usage does not reflect it. Second, the all-degeminating (45c) is in the third place in the usage, but is supposed to be dispreferred in the formal analysis. The usage here, however, depends on Google hit counts, and it may not be reliable: there may be typographical mistakes, some of the usage could be largely depending on one influential source with anomaly, and Japanese Internet slangs are also known for its tendency to degeminate (Uchiyama, 2010). In section 6, I conduct a nonce-word adaptation experiment to overcome these problems.

## 5.4 Summary and Predictions

The following Hasse diagram summarizes the configuration for the lexical stratum discussed so far.

(46) The constraints and relative rankings



One of the prediction this configuration makes is that it is always the categorically positionally faithful candidate that wins. Thus, when Japanese adapts words such as *cappelletti* (a type of pasta), where there are both liquid and non-liquid geminates in weak position, the winner according to this configuration is [ka.pe.réT.ti], and this is actually what we usually see in usage. Unfortunately, the usage seems to have fossilized as [ka.pe.réT.ti], and Google search does not tell us about further variation. However, assuming the constraints and ranking as in (46), the second-best output for an input of the form /cvC.cvR.cvC.cv/ are [cv.cv:cvC.cv] and [cv:cv.cvC.cv].

Assuming that the rankings among faithfulness constraints are fixed, moving around the markedness constraints NOGEM[R] and NOGEM would predict different outputs as to the degree of nativization. For example, *azzurri* ‘blue-PL’ is most commonly adapted as [a.zu:ri], and ranking NOGEM[R] in between MAX-'σ[μ] and IDENT-'σ[μ]

would predict degemination of the obstruent geminates in the weak position and compensatory lengthening of the liquid geminates in the strong position.

(47) Degemination and vowel lengthening in *azzurri* with Google hits

	/aZ.zúR.ri/	<i>MAX</i> -' $\sigma$ [ $\mu$ ]	<i>NOGEM</i> [R]	<i>IDENT</i> -' $\sigma$ [ $\mu$ ]	<i>NOGEM</i>	<i>IDENT</i> [ $\mu$ ]	<i>MAX</i> [ $\mu$ ]	Google Hits
a.	a.zu:ri			**		**		364,000
b.	aZ.zu:ri			**	*!	**		46,800
c.	a.zu.ri	*!		*		**	**	17,300
d.	a.zuR.ri		*!		*	*	*	405
e.	a:zu.ri	*!		*		***	*	353
f.	aZ.zuR.ri		*!		**			307
g.	aZ.zu.ri	*!		*	*	*	*	249
h.	a:zu:ri			**		***!*		4
i.	a:zuR.ri		*!		*	**		0

## 6 A Nonce-adaptation Survey

The Google hit search, while giving us a rough idea on which forms are more frequently used, includes noises such as typos or homographs. In order to get closer to the actual preference of Japanese speakers as to the position and types of geminates in loanwords, I conducted an online nonce-adaptation rating questionnaire. In this survey, native speakers of Japanese were exposed to Italian nonce-words represented in Italian orthography, and were asked to rate them on a scale of 1 (dispreferred) to 10 (preferred).

### 6.1 Methods

#### 6.1.1 Participants

Participants were 27 native speakers of Japanese, recruited through personal connection. Demographically, the participants are thought to be mainly between 20 to 30 years old males and females. They were expected to have some proficiency in English and Roman alphabets, but not as much familiarity with Italian orthography.<sup>21</sup> However, given that the Italian writing system is quite transparent and the representation of geminates in Italian orthography is about the same as how Roman alphabets are used to transcribe Japanese geminates (i.e. repetition of the same letter) in roma-ji (Roman-alphabet transcription of Japanese), I expect not much difficulty for the participants in interpreting the presence of geminates. Orthographically opaque geminates (‘sci’ for [ʃʃ] or ‘gl’ [ʎʎ]) were excluded from the stimuli.

Participants were assumed to be literate in Japanese writing system as well, where the presence of an obstruent geminate is indicated with ‘っ’. The presence of a long

---

<sup>21</sup>In the questionnaire, all except one reported that they have studied English, and the one who did not report English had studied German and Italian.

vowel is also straightforward, indicated by ‘—’ immediately following the syllabary for vowel.

The survey results were collected anonymously, and anyone who has access to the internet was able to take it. 31 people took the survey, of which four were not included in the analysis as they gave very low rating across the board, indicating some floor effect or miscommunication of the task.

### 6.1.2 Material

The primary objective of this survey was to see the effect of position and of the type of geminates on degemination. The stimuli consisted of pairs of a trisyllabic nonce-word in Italian (the *input*) and its possible adaptation forms (the *output* candidates). The context for adaptation was in a restaurant, where the menu only indicated the name of the dishes in Italian orthography (it is rare in Japan that there is no Japanese orthography accompanying it, but I gave the context to make the task easier for participants). The participants were not told specifically that it was an Italian restaurant.

The tokens were controlled for word-length (three syllables) and number of geminates (two in each token, in the antepenultimate and penultimate syllable<sup>22</sup>), and stress position. Italian stress for these nonce-words were assumed to be in the penult, given that a heavy penult attracts stress in Italian, and loanwords in Japanese often receive accent on the antepenultimate mora. In other words, the survey assumed that there would be no divergence of the accent position between the Italian parsing and the Japanese parsing, the Japanese pitch accent being LHL. The survey was not designed to answer the question as to which of the accent, input or output, is responsible for the preservation of a geminate. Rather, it focused on recreating the effect of position. My analysis

---

<sup>22</sup>Geminates in the ultimate syllable is not possible, as both Italian and Japanese do not allow lexical word-final geminates.

predicted that outputs with positional degemination would be preferred over outputs with anti-degemination (i.e. degeminating a penultimate geminate and preserving an antepenultimate geminate).

As for the effect of the types of geminates, I considered two types, voiceless obstruent and liquid geminates. From my analysis, I expected that liquid geminates would be more prone to degemination or compensatory lengthening (see Kawahara 2007 on the rapport between the sonority hierarchy and the well-formedness of geminates).

Given these premises, the tokens included four types of input presented in (48). For each type, there were three input tokens. The two types of geminates were represented as ‘g’ or ‘G’ (voiceless geminates) and ‘r’ or ‘R’ (liquid geminates); lower case indicating that they were in *weak* position (antepenultima) and upper case indicating that they are in *strong* position (second geminate). Light syllables will be indicated as ‘l’ or ‘L’.

(48) Italian nonce-words (input)

type	tokens		
gGl	ciuffocco	doffoccio	bottossa
gRl	eppella	ducciolla	tuttullu
rGl	gorruppa	vorrotto	forrotto
rRl	collerre	ciollerre	billorro

For each input token, there were 5 output candidates, narrowed down from the logically possible 9 candidates. The 9 logical possibilities were permutations of possible operations on each input geminate: geminates could be preserved (*faith*), undergo compensatory lengthening (*comp*), or be degeminated (*degem*). (49) shows all 9 possible candidates for an input gGL, *bottossa*, as an example. ‘v’ and ‘V’ indicate compensatory lengthening. The rightmost two columns indicate the operation each geminate

went through.

(49) The logically possible 9 output forms

	output	example	weak gem	strong gem
1.	gGl	boT.toS.sa	faith	faith
2.	gVl	boT.to:sa	faith	comp
3.	gLl	boT.to.sa	faith	degem
4.	vGl	bo:.toS.sa	comp	faith
5.	vVl	bo:.to:sa	comp	comp
6.	vLl	bo:.to.sa	comp	degem
7.	lGl	bo.toS.sa	degem	faith
8.	lVl	bo.to:sa	degem	comp
9.	lLl	bo.to.sa	degem	degem

Among these 9 adaptation possibilities, the candidates that participants were given in the rating task were the following (relabelled as A, B, C, D, E here):

(50) Types of rated output forms

	output	example	weak gem	strong gem
A.	lGl	bo.toS.sa	degem	faith
B.	vGl	bo:.toS.sa	comp	faith
C.	lVl	bo.to:sa	degem	comp
D.	gVl	boT.to:sa	faith	comp
E.	gLl	boT.to.sa	faith	degem

The candidate inventory purposefully did not contain the fully-faithful option, with two geminates in the output. This was to avoid a bimodal distribution of the rating results, in which the participants rate the fully-faithful candidate very high and all the

other ones very low, assuming that the first stage of the adaptation would be mostly faithful. Geminating candidates were also not considered.

The critical tokens were generated by assigning consonants and vowels to each position using the randomizing function on Microsoft Office Excel. For the critical tokens, there were total of 60 trials. There were as many fillers as the critical tokens, making the total number of trials 120. The stimuli were presented using Google Forms, with two practice trials. There were 5 blocks, and at the beginning of each block, the participants saw the instruction. They were encouraged to take as much break as they want between the third and fourth blocks.

The 60 fillers were also trisyllabic Italian nonce-words, but without any geminates. 50 of them were taken from Colombo (1992), and the rest were supplemented from Zoccolotti et al. (2005).

(51) Fillers

batilo, bildese, birfola, birtona, blosidi, boltici, bortaca, borteso, bortume, canfrosto, cegape, celimo, cirtora, dilone, dinuro, drivule, fanziane, fastanda, flenesta, fromile, grocelso, iselo, laromo, linebre, lintere, livero, loraia, marlipo, meribe, mevino, olina, onfilì, ostura, panchefa, pifato, pirtoci, polaso, poracca, potide, prigiosa, primosta, ravele, rebolo, rudomi, rulate, sintuce, stebore, stevono, stilega, storubo, strebafè, strotula, svepano, tegresto, tirloni, trofulo, trolica, tuposo, vielota, virpico, zerlido

### 6.1.3 Procedure

Participants were exposed to pairs of input and output, and were asked to rate each output using radio buttons, with a scale of 1 to 10, 1 being “not likely to say”, and 10 being “likely to say”. (52) is an illustration of the interface for input *eppella* and output



[eppe:ra]. Parts of the survey are also presented in the Appendix.

(52) *eppella*

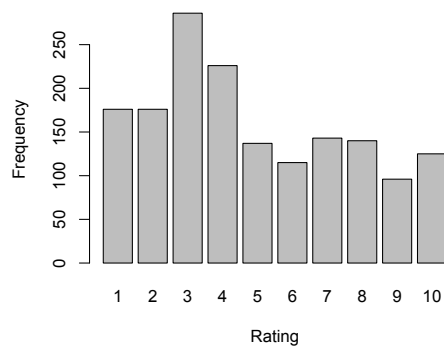
[eppe:ra]            1   2   3   4   5   6   7   8   9   10  
unlikely to say    ○   ○   ○   ○   ○   ○   ○   ○   ○   ○   ○   likely to say

The order of questions in each section was randomized using the randomizing feature of Google Form. Each section contained one instance of one input nonce-word, and which of the adaptation candidate forms shows up in each section was randomly decided using Excel.

## 6.2 Results and Discussion

I obtained 1620 responses for the critical tokens. The responses were analyzed using the statistical programming environment R (R Core Team, 2013) and *lme4* (Bates et al., 2012). The mean of all responses was 4.88, and as can be seen in Figure 2, the rating responses were not normally distributed – participants did not like the suggested output in general, the most popular responses being 3 and 4 out of 10. This is perhaps because participants were not provided with the most faithful output, preserving geminates.

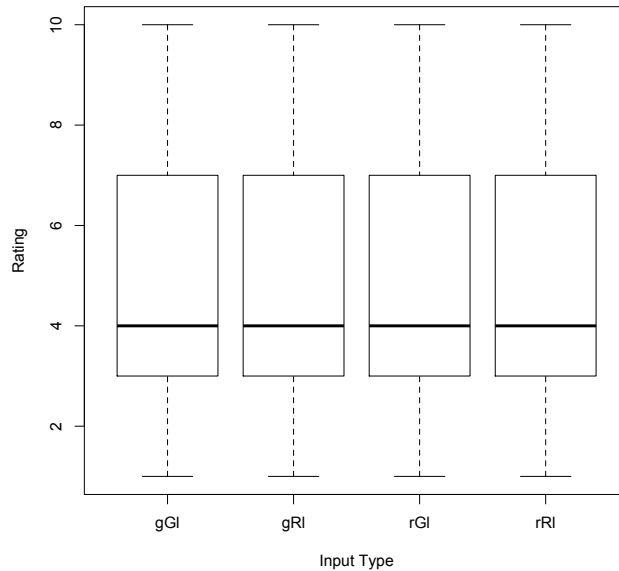
Figure 2: Frequency of rating responses



### 6.2.1 Effect of Types of Gemimates

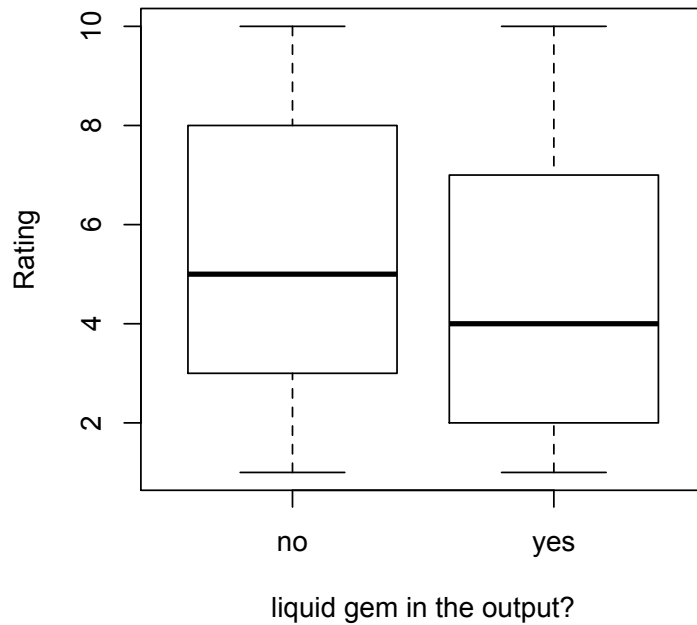
At the baseline, the types and combination of gemimates in the input did not affect the responses ( $\chi^2(1) = 0.60$ , n.s.). The average ratings for each types of input in Figure 3 are quite even.

Figure 3: Average ratings for each type of input



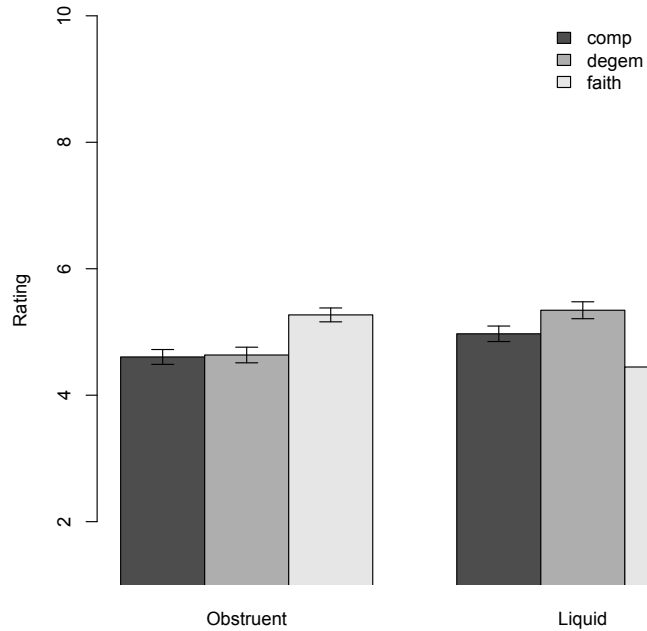
However, the types of gemimates that appear in the output had significant effect on the rating responses. There was a preference to outputs that do not contain liquid gemimates, as can be seen in Figure 4. Outputs containing liquid gemimates were rated lower than outputs without one (this category includes outputs which has no gemimates at all, as a result of degemination and compensatory lengthening). I performed a linear mixed effects analysis of the relationship between the rating and the presence of a liquid geminate in the output. As random effects, I had intercepts for participants and words. For the fixed effects, I only had the presence/absence of a liquid geminate in the output. The result was statistically significant ( $\chi^2(1) = 35.41$ ,  $p < 0.001$ ).

Figure 4: Average ratings for outputs containing liquid geminates (“yes”) and outputs without one (“no”).



Furthermore, different types of geminates preferred different operations to undergo, as is evident from Figure 5. Through a linear mixed effects analysis with intercepts for participants and words as random effects, and an ANOVA, the interaction of the type of geminates and operations was confirmed ( $\chi^2(1) = 63.77, p < 0.001$ ). In general, obstruent geminates are preferred to be preserved, while liquid geminates are preferred to be degeminated. The second choice for liquid geminates is compensatory lengthening rather than preservation. While it is a much more subtle tendency, obstruent geminates seem to prefer degemination rather than compensatory lengthening.

Figure 5: Preferences for operation per types of geminates. Note that the same rating is used twice to calculate this, once for weak geminates, once again for strong geminates.



## 6.2.2 Effect of Position of Geminates

First, by a simple comparison between outputs that include geminates in the weak position and outputs that include geminates in the strong position (Figure 6), it is evident that the latter got higher ratings ( $\chi^2(1) = 22.26, p < 0.001$ ).

In terms of the operation on all input geminates, from visual inspection of Figure 7a and 7b, there is a trend such that geminates in the weak position are preferred to be degeminated or compensatory lengthened, rather than preserved; geminates in the strong position are preferred to be preserved. The effect, however, turned out to be significant for weak geminates ( $\chi^2(1) = 20.98, p < 0.001$ ), but not for strong geminates ( $\chi^2(1) = 0.00, p < 1$ ).

Figure 6: Average Rating for outputs including geminates in either weak position (C1) or strong position (C2) (outputs without any geminates (i.e. no preservation) were eliminated from the analysis).

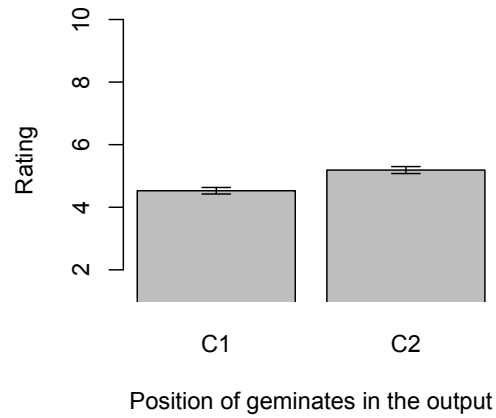
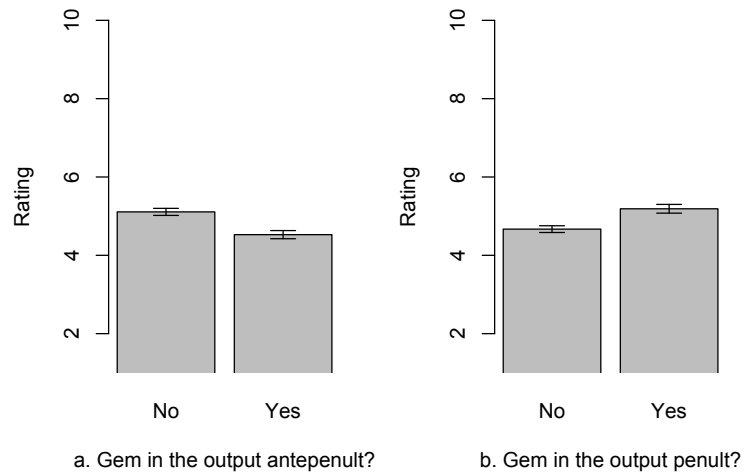


Figure 7: Average ratings for outputs depending on whether they contained a geminate, (a) in the weak position or (b) strong position.



This might be due to my analysis method that is treating degemination and compen-

satory lengthening bundled together. Graphs that go a bit further in details are Figure 8a and 8b.

Figure 8: Preferred operation for geminates in (a) weak position and (b) strong position.

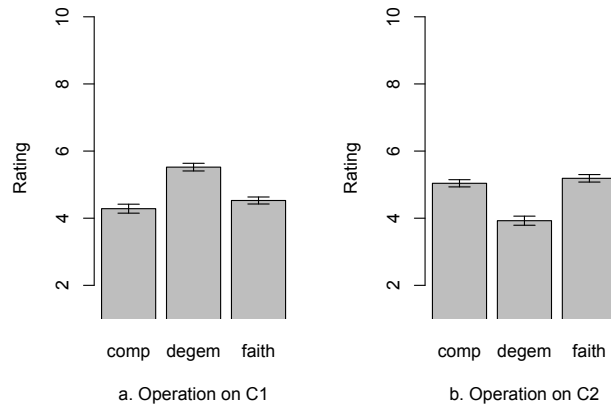
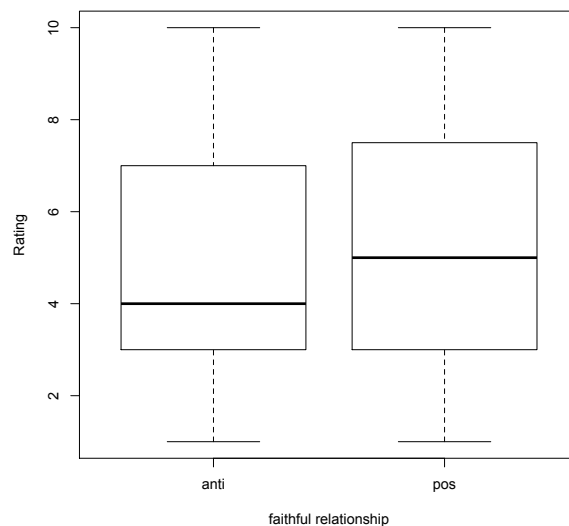


Figure 8a and 8b represent an expected, clear contrast between the two positions. First, weak geminates are preferred to be degeminated, while degemination is the least favorite operation for strong geminates. Meanwhile, the difference between the preference for compensatory lengthening and preservation is not as clear. In both positions, it seems like preservation is slightly favored over compensatory lengthening, but the effect is visually subtle. What is less murky is that compensatory lengthening and preservation are more preferred in the strong position than in weak position, and this tendency conforms to our prediction that strong positions require more faithfulness than weak positions do. While the tendency generally points to the prediction of our formal analysis so far, not everything turned out to be statistically significant: as results of an ANOVA over linear mixed effects models, the effect of the operation over weak position on the rating turned out to be significant ( $\chi^2(1) = 75.08, p < 0.001$ ), while non-significant for strong position ( $\chi^2(1) = 0.00, p < 1$ ). The additive effect of the operation on weak position and strong position, on the other hand, turned out to be significant

( $\chi^2(1) = 75.96, p < 0.001$ ).

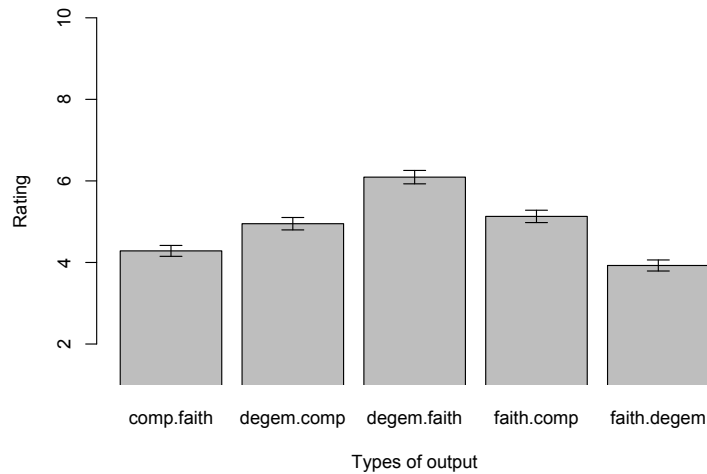
Another perspective to take regarding the positional effect, and one of the aims of this experimental design, is the relationship between geminates in the strong and weak position. My formal analysis predicts that forms preserving a weak geminate while degeminating a strong geminate are not a viable option – or at least dispreferred. This is due to the fact that in such forms, the degree of faithfulness is lower in the privileged position, than in a non-privileged position. Figure 9 compares the ratings of two categories of output types: “anti” and “pos”, where “anti” refers to outputs in which faithfulness is lower in strong position, and “pos” refers to outputs in which faithfulness is higher in strong position. There was no outputs candidates that embodied the same amount of faithfulness in the two positions in this experiment. The effect was confirmed to be significant through a linear mixed effects model analysis and an ANOVA ( $\chi^2(1) = 20.98, p < 0.001$ ).

Figure 9: Preferences depending on the faithfulness relationship between strong and weak geminates.



Finally, Figure 10 gives us an overall picture of the positional effect, showing the preference for each combination of operations in the two positions. Another ANOVA confirmed that the influence of these combination of operations, regardless of the types of geminates, is significant ( $\chi^2(1) = 151.04, p < 0.001$ ).

Figure 10: Preferences depending on the output types.



The table below lays out the same hierarchy using our schema presented in (50):

(53)

status	degem.fai th	>	fai th.comp	>	degem.comp	>	comp.fai th	>	fai th.degem
types	lGl		gVl		lVl		vGl		gLl
example	bo.toS.sa		boT.to:sa		bo.to:sa		bo:toS.sa		boT.to.sa

There is a clear preference for outputs that preserve the strong geminate while degeminating the weak geminate, as predicted in our analysis. The low rating for outputs that preserve the weak geminate while degeminating the strong geminate is also predicted. It is interesting, though, that the preference for outputs that maintain more faithfulness in strong position than in weak position does not show up in this hierarchy: preserving a weak geminate while compensatory lengthening a strong geminate (gVl) is more

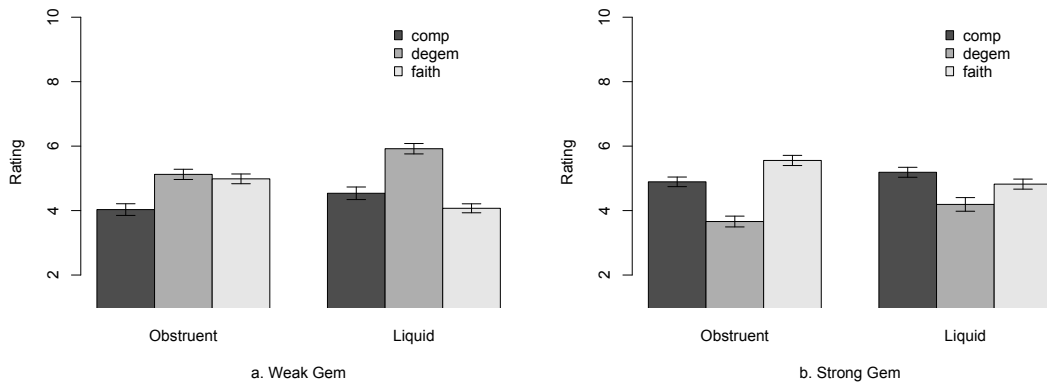


popular than the opposite pattern (vGI). As a possible reason, it might be that the rating was very high for the pairs of gRI (input) and gVI (output). That is, high preference for strong liquid geminates undergoing compensatory lengthening.

### 6.2.3 Interaction of Position and Types of Geminates

The hierarchy of outputs in Figure 10 seems to be pointing to some interactions between positions and types of geminates in the stimuli. Figure 11a and 11b allow us to browse through them.

Figure 11: Preferences for operation per types of geminates within (a) weak geminates and (b) strong geminates.



First, comparing the operations on obstruent geminates in Figure 11a and 11b, we notice that preservation of obstruent geminates is more popular when they are in strong positions. And more robustly, degemination of strong obstruent geminates is much less preferred compared to degemination of weak obstruent geminates. Furthermore, compensatory lengthening of obstruent geminates is more preferred in strong position than in weak position.

As for liquid geminates, degemination in weak position is largely preferred, while the choice is not as popular in strong position. It is also striking that the preservation of strong liquid geminates is allowed in strong position. This is actually in line with the formal analysis provided in (45), whereby there was a mismatch between the prediction of the analysis and the informal indicator of usage. Liquid geminates in strong position are largely allowed. In addition, the high preference for strong liquid geminates addresses the earlier question about the gVl pattern being more popular than vGl, although vGl is preserving more in the strong position: in fact, compensatory lengthening is most preferred as operation for strong liquid geminates. Note also that compensatory lengthening is usually more preferred for liquid geminates than for obstruent geminates. This may connect to the sonorancy of geminates discussed in Kawahara (2005). (54) is a summary of which operation is preferred for each type and position.

(54) Summary of operation preferences

Type	Position	Operations
obstruent	weak	degem > faith > comp
	strong	faith > comp > degem
liquid	weak	degem > comp > faith
	strong	comp > faith > degem

This result has interesting correspondence with (43). The preference shown in (43) for liquid geminates correspond to the preference for obstruents in (54).

The effect of positions and types was assessed as follows: I performed a linear mixed effects analysis of the relationship among the geminate types in different positions. I constructed models with possible combinations of the four factors illustrated in (55) (random effects were the intercepts for subject and word) and ran an ANOVA to see which model has the best fit.

(55) Models (shaded cells had significant effect)

Fixed Effects	Chisq	p value
a. weak gem type	0.19	n.s.
b. strong gem type	0.12	p < 0.001
c. op on weak gem	74.78	p < 0.001
d. operation on strong gem	0.00	n.s.
e. weak gem type + op on weak gem	14.96	p < 0.001
f. strong gem type + op on strong gem	0.00	n.s.
g. weak gem type + op on strong gem	0.00	n.s.
h. strong gem type + op on weak gem	14.88	p < 0.001
i. weak gem type + op on weak gem + strong gem type + op on strong gem	62.92	p < 0.001
j. weak gem type * op on weak gem	43.29	p < 0.001
k. strong gem type * op on strong gem	0.00	n.s.
l. weak gem type * op on strong gem	18.69	p < 0.001
m. strong gem type * op on weak gem	0.00	n.s.
n. weak gem type * op on weak gem * strong gem type * op on strong gem	95.15	p < 0.001

The best model seems to be (55n), the one with the interaction of all four factors (types of geminates in weak position and in strong position), accounting for about 95% of the data. The additive effect of the four factors in (55i) is also good, but accounts for slightly less than (55n).

However, the operation on weak geminates already accounts for more than 70% alone – another ANOVA showed that (55n, l) are significantly different from (55c) ( $\chi^2(1) = 76.51$ ,  $p < 0.001$ ;  $\chi^2(1) = 95.15$ ,  $p < 0.001$ ), but it is interesting that the operation on weak geminates has such a great impact on the whole rating responses. It seems to underline the importance of degeminating weak geminates, an effect evident in Figure 8. The other significant effect among the individual effects is the type of strong geminates, (55b). This is indicative of a robust distinction between obstruent

and liquid geminates in strong position, whereby obstruents are preserved and liquids are degeminated or undergo compensatory lengthening. The additive effect in (55e) turned out not to be significantly different from (55c) ( $\chi^2(1) = 0.32, p < 1$ ), as well as (55h) ( $\chi^2(1) = 0.19, p < 1$ ), and it is assumed that it comes from the effect of the operation on weak geminates. The same applies for (55j) ( $\chi^2(1) = 0.32, p < 1$ ): (55c) seems to be responsible for the interaction of the type and operation of weak geminate.

On the other hand, for (55l), the interaction of the two individually non-significant effects, the type of weak geminate and the operation on the strong geminates, seems to be real.

### **6.3 Summary and Further Development**

The general tendencies observed in the survey mostly conform to the prediction of the formal analysis, in that geminates in penultimate syllables are preferred to be kept if it is an obstruent, and compensatory lengthened if it is a liquid; whereas geminates in antepenultimate syllables are preferred to be degeminated in general, but more for liquid geminates than for obstruent geminates. However, some of the results including the hierarchy in (53) could not be predicted from the formal analysis, and there are several ways in which the survey could be developed further. This subsection summarizes the points of improvement for this survey.

First, some of the orthographic property of the input may have influenced the rating responses: inputs including “ciu” and “rr” probably should have been avoided, as “ciu” to [tʃiu] does not exist in the romaji system, and “rr” sometimes undergoes vowel epenthesis instead of being adapted as geminate. Therefore, eliminating these segments would have made sure the data is clean.

Meanwhile, a more radical remedy would be to develop this survey to an auditory

perceptual one, which not only will overcome these shortcomings, but also would make this experiment more compatible with my core claim about the privileged position for geminate preservation being the Italian stressed syllable, rather than the accented syllable in Japanese. In addition, it may answer the question about the sonority being in play – in the current survey, there is no guarantee that participants were aware of pronunciation of the unfamiliar segments.

Finally, rating using Likert scale may not be the ideal task, as it relies on meta-linguistic judgments. While it would still be a similar kind of experiment, a pairwise comparison of output may be more appropriate (one downside of pairwise comparison is that it would make the experiment much longer).

## **7 Conclusion**

This paper developed an analysis of positional degemination in Japanese loanword from Italian. I claimed that the positional effect on degemination in Japanese loanwords from Italian can be explained as positional neutralization based on Italian stress position (which can ultimately be overwritten with Japanese accentual pattern). The effect can be formalized using the positional faithfulness schema, assuming an Output-output correspondence relationship between the source and adapted forms.

The analysis predicts that, when the positional constraint is effective, the optimal adaptation would be the one that preserves the strong geminate, degeminating the weak geminate; it also predicts that the opposite pattern is an impossible adaptation. My nonce-word adaptation survey confirmed the tendency, and showed the robust effect of position. The positional effect can even override the segmental restrictions.

The survey, however, did not address the question as to whether the prominence-based positional faithfulness is based on the Italian stress or Japanese accent. In order

to address that, it would be useful to develop a perceptual experiment of similar design, but with varying Italian stress position.

The analysis took Tanaka's (2007) generalization further to identify the privileged position as the syllable bearing stress in Italian, which most of the time ends up in the penultimate. This generalization renders Tanaka's (2007) observation a rather trivial one. It did not address, however, the tendency that antepenultimate geminates also have relatively high rate of preservation. It would be necessary to look into forms such as *glissándo* → [gu.riS.san.do], *crescendo* → [ku.reS.fen.do], or *toccáta* → [toK.kar.ta], where non-stressed antepenultimate geminates are preserved. Furthermore, some longer forms such as *cappuccino* → [kaP.pu.tʃi.no] would have to be accounted for, in conjunction with the word-length effect. Alternative analyses should perhaps include the prosody of Japanese, thus dealing with more markedness constraints.

Finally, on the theoretical side, the phenomena discussed in this paper would also provide insights in frameworks other than standard OT, such as Harmonic Grammar (McCarvel and Kaplan, 2013) in terms of constraint interaction, or Harmonic Serialism (McCarthy, 2000), departing from parallel OT.

# Appendix

## The Survey

**MENU**

あなたはレストランに来ています。メニューの冊子を開くと、料理の名前がアルファベットで書かれています。注文するとき、色々な読み方がありますが、あなたはどのように読みますか？

You are in a restaurant. In the menu, you find the names of the dishes written in alphabet. How do you think you would pronounce them? There are various possible pronunciation of these names.

**説明 Explanation**

与えられた料理名 (アルファベット) に対し、その下に書かれたカタカナについて、レストランで言うと思う度合いを評価してください。試しに口に出してみてもかまいませんが、あまり深く考えずに直感で評価してください。

評価は1 (言うと思わない)~10 (言うと思う)の10段階です。

次のページで、練習してみましょう。

For the pronunciation given in katakana for each name of the dish (given in alphabet), please respond the rating of how likely you would say it. You can try actually saying them, but do not ponder too much, and please respond intuitively. The rating is from 1 (least likely to say) to 10 (most likely to say). Let's practice in the following page.

Continue » 12% completed

Powered by Google Forms

**MENU**

\* Required

**練習 Practice**

与えられた料理名 (アルファベット) に対し、その下に書かれたカタカナについて、レストランで言うと思う度合いを評価してください。試しに口に出してみてもかまいませんが、あまり深く考えずに直感で評価してください。

評価は1 (言うと思わない)~10 (言うと思う)の10段階です。

For the pronunciation given in katakana for each name of the dish (given in alphabet), please respond the rating of how likely you would say it. You can try actually saying them, but do not ponder too much, and please respond intuitively. The rating is from 1 (least likely to say) to 10 (most likely to say).

**gelato \***  
グレート

1 2 3 4 5 6 7 8 9 10

I don't think I'll say           I think I'll say

**spaghetti \***  
スパゲッティ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

それでは、"Continue"をクリックして本番を始めてください。

時間制限はありません。質問は全部で5セクションあります。フルコースなので数が多いですが、食休みをとりつつ頑張ってください。

Now, click on the "Continue" button to start the survey. There is no time limit. There are 5 sections. This is a lot of questions as it is a full-course meal, but I hope you'll make it through with some intervals in-between sections.

Back Continue » 25% completed

**MENU**

\* Required

**1. 前菜 Appetizers**

それではよいよ、本番です。Let's start then. (the instruction repeated)

与えられた料理名 (アルファベット) に対し、その下に書かれたカタカナを、レストランで言うと思う度合いで評価してください。試しに口に出してみてもかまいませんが、あまり深く考えずに直感で評価してください。

評価は1 (言うと思わない)~10 (言うと思う)の10段階です。

**ruiate \***  
ルラテ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**hatilo \***  
ハティロ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**celimo \***  
チェリモ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**doffocio \***  
ドフォーチオ

1 2 3 4 5 6 7 8 9 10

**hottouss \***  
ポートルッソ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**hillorro \***  
ビロロロ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**meviao \***  
メビオ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**trolica \***  
テロリカ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

**strotala \***  
ステロタラ

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

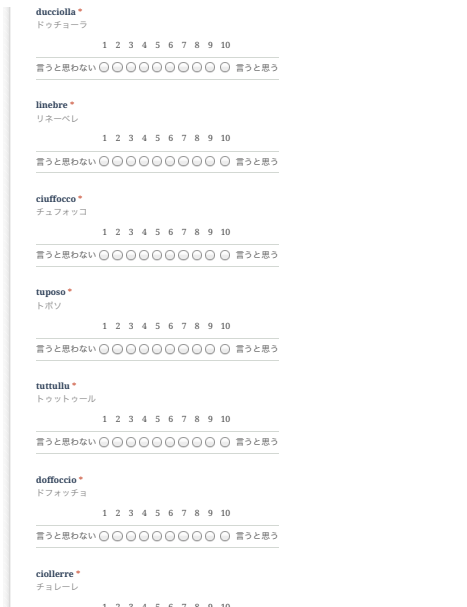
**forrotto \***  
フォーロット

1 2 3 4 5 6 7 8 9 10

言うと思わない           言うと思う

Back Continue » 37% completed

Powered by Google Forms





## References

- Anttila, Arto. 1997. Deriving variation from grammar. *Variation, Change and Phonological Theory* 146:35.
- Bates, Douglas, Martin Maechler, and Ben Bolker. 2012. *lme4: Linear mixed-effects models using Eigen and syntax*.
- Beckman, Jill N. 1999. *Positional Faithfulness: An optimality theoretic treatment of phonological asymmetries*. New York and London: Garland Publishing, Inc.
- Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.
- Brennan, Jon. 2006. A weight problem: Vowel length in classical latin. Handout for presentation at the 8th SUNY/CUNY/NYU Miniconference.
- Colombo, Lucia. 1992. Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance* 18:987.
- D’Imperio, Mariapaola, and Sam Rosenthal. 1999. Phonetics and phonology of main stress in Italian. *Phonology* 16:1–28.
- Esposito, Anna, and Maria Gabriella Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America* 106:2051–2062.
- Irwin, Mark. 2011. *Loanwords in Japanese*. John Benjamins Publishing.
- Itô, Junko, and Armin Mester. 1993. Japanese phonology: constraint domains and structure preservation. In *A Handbook of Phonological Theory*, Blackwell Handbooks in Linguistics Series. Citeseer.
- Itô, Junko, and Armin Mester. 1995a. The core-periphery structure of the lexicon and constraints on reranking. In *Papers in Optimality Theory*, ed. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh, 181–210. Amherst, Mass.: GLSA Publications.
- Itô, Junko, and Armin Mester. 1995b. *The Handbook of Phonological Theory*, chapter Japanese Phonology, 817–838. Blackwell.
- Itô, Junko, and Armin Mester. 1999a. The phonological lexicon. In *The Handbook of Japanese Linguistics*, ed. Natsuko Tsujimura, 62–100. Oxford: Blackwell.
- Itô, Junko, and Armin Mester. 1999b. Realignment. *The Prosody-Morphology Interface* 188–217.

- Itô, Junko, and Armin Mester. 2003. *Japanese Morphophonemics: Markedness and word structure*. MIT Press.
- Itô, Junko, and Armin Mester. 2014. Unaccentedness in Japanese. To appear, November 2014.
- Katayama, Motoko. 1998. Optimality theory and Japanese loanword phonology. Ph.d., Department of Linguistics, University of California at Santa Cruz.
- Kawagoe, Itsue. to appear. The phonology of *sokuon*, or geminate obstruents.
- Kawagoe, Itsue, and Masako Arai. 2002. Syakuyoogo ni okeru sokuon [Consonant gemination in loanwords]. *Onsei Kenkyu* 6:53–66.
- Kawahara, Shigeto. 2005. Sonorancy and geminacy .
- Kawahara, Shigeto. to appear. The phonetics of obstruent geminates, sokuon. *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology* .
- Kenstowicz, Michael, and Atiwong Suchato. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116:921–949.
- Kiparsky, Paul. 1968. How abstract is phonology? In *Three Dimensions of Linguistic Theory*, ed. Osamu Fujimura, 5–56. Tokyo: TEC.
- Krämer, Martin. 2009. *The Phonology of Italian*. OUP Oxford.
- Kubozono, Haruo. 1995. *Gokeisei to on-in-kouzou [Word formation and phonological structure]*. Tokyo: Kuroshio.
- Kubozono, Haruo. 2006. Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua* 116:1140–1170.
- Kubozono, Haruo, Junko Itô, and Armin Mester. 2009. Consonant gemination in Japanese loanword phonology. In *Current Issues in Unity and Diversity of Languages. Collection of Papers Selected from the 18th International Congress of Linguists [CIL 18]*, ed. The Linguistic Society of Korea, 953–973. Republic of Korea: Dongam Publishing Co.
- Labrune, Laurence. 2012. *The Phonology of Japanese*. Oxford University Press.
- Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. *Harmonic Grammar - a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations*. Citeseer.
- Lovins, Julie B. 1975. *Loanwords and the Phonological Structure of Japanese*. IULC.

- Martin, Samuel E. 1952. Morphophonemics of standard colloquial Japanese. *Language* 7–115.
- McCarthy, John J. 2000. *Harmonic Serialism and Parallelism*.
- McCarthy, John J, and Alan Prince. 1995. *Faithfulness and Reduplicative Identity*.
- McCarvel, Miranda, and Aaron Kaplan. 2013. Positional faithfulness in harmonic grammar. *Phonology* 2013.
- McCawley, James D. 1968. *The Phonological Component of a Grammar of Japanese*. 2. Mouton.
- Ministry of Education, Culture, Sports, Science, and Technology. 2008. *Shogakko gakushu shido yoryo kaisetsu kokugo hen [Explanation of the course of study for elementary school concerning Japanese]*. Tokyo.
- Morén, Bruce. 2001. *Distinctiveness, Coercion and Sonority: A unified theory of weight*. Psychology Press.
- NHK Hōsō Bunka Kenkyūjo, ed. 1998. *NHK Nihongo hatsuon akusento jiten [NHK Japanese pronunciation accent dictionary]*. Nihon Hōsō Shuppan Kyokai.
- Nishimura, Kohei. 2003. Lyman's Law in Loanwords. Master's thesis, Nagoya University.
- Nishimura, Kohei. 2006. Lyman's Law in Loanwords. *Phonological Studies* 9:83–90.
- Ohye, Saburo. 1967. Gairaigo chū no sokuon ni kansuru ichi kōsatsu [The mora phoneme /Q/ in English loanwords in Japanese]. In *Onsei no Kenkyu [Bulletin of Japan Phonetics Society]*, volume 13, 111–121.
- Payne, Elinor M. 2005. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35:153–181.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Revithiadou, Anthoula. 1999. Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems. Doctoral Dissertation.
- Sano, Shin-ichiro, and Shigeto Kawahara. 2013. A corpus-based study of geminate devoicing in Japanese: The role of the ocp and external factors. *Gengo Kenkyu [Journal of the Linguistic Society of Japan]* 144:103–118.

- Sanseidō Henshūjo, ed. 2010. *Konsaisu katakana go jiten [Concise Katakana-word dictionary]*. Tokyo: Sanseidō, 4th edition.
- Sanseidō Henshūjo, ed. 2015. *Sūpā daijirin 3.0 web kōshinban*. Tokyo: Sanseidō.
- Sanseidō Henshūjo, and Hideo Satake, ed. 2010. *Deirī konsaisu kokugo jiten [Daily concise Japanese dictionary]*. Tokyo: Sanseidō, 5th edition.
- Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge University Press.
- Shinmura, Izuru. 1998. *Kōjien*. Tokyo: Iwanami Shoten, 5th edition.
- Smith, Jennifer L. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. *Japanese/Korean Linguistics* 14:63–74.
- Tanaka, Shin'ichi. 2007. Itariago no juushiin to sokuon keisei [Geminate consonants in Italian and sokuon in Japanese]. *Proceedings of the 134th Meeting of the Linguistic Society of Japan* 252–257.
- Tanaka, Shin'ichi, and Haruo Kubozono. 2008. Itariago no nijuu-shiin ni taisuru sokuon chikaku [Japanese listeners' perception of Italian geminate consonants]. In *Proceedings of the 22th Annual Meeting of the Phonetic Society of Japan*, 61–66.
- Tanioka, Takeo. 1998. *Konsaisu gaikoku chimei jiten [Concise foreign place name dictionary]*. Tokyo: Sanseidō, 3rd edition.
- Uchiyama, Hiroshi. 2010. Netto no nihon go: 2 channeru to nikoniko dogo o chushin ni [An analysis of Japanese used in the net world: 2channel and NicoNikoDoga]. *Chiiki Seisaku Kagaku Kenkyu* 7:219–236.
- Vance, Timothy J. 1997. *An Introduction to Japanese Phonology*. SUNY Press.
- Vance, Timothy J. 2008. *The Sounds of Japanese with Audio CD*. Cambridge University Press.
- Yamada, Tadao, Takeru Shibata, Kenji Sakai, Yasuo Kuramochi, Akio Yamada, Zendō Uwano, Ijima Masahiro, and Sasahara Hiroyuki, ed. 2011. *Shinmeikai kokugo jiten*. Tokyo: Sanseidō, 7th edition.
- Yip, Moira. 2002. Perceptual influences in Cantonese loanword phonology. *Journal of the Phonetic Society of Japan* 6:4–21.
- Zanichelli Editore Spa, ed. 2013. *Lo Zingarelli 2013*. Zanichelli.
- Zoccolotti, Pierluigi, Maria De Luca, Gloria Di Filippo, Anna Judica, and Donatella Spinelli. 2005. Prova di lettura di parole e non parole. *IRCCS Fondazione Santa Lucia*.