Gemination in Japanese Loanwords

Derek Andersen

Department of Linguistics

Stony Brook University

December 10, 2019

1   Introduction

Japanese borrowings of English (and other) words provide a rich source of data for the analysis of the gemination of certain consonantal segments in Japanese, and in particular, the types of segments for which Japanese phonology allows gemination, and the phonological implications of gemination with respect to both the input (source word) and output (phonemicized) forms. My analysis of gemination in Japanese loanword phonology will be focusing mainly on the *voiced obstruent* geminates /bb/, /dd/, /gg/, and the motivation for gemination in loanwords. Specifically, the issues that arise with foreign word gemination when compared to the native phonotactics of Japanese. In section 2, I present some background on the native phonemic inventory of Japanese, and some evidence for the restriction on voiced geminates in the native phonology. In section 3, I give an overview of the distribution of the lexical strata of Japanese. In section 4, I discuss some cases of gemination in the creation of Japanese loanwords, and present an Optimality Theory (Prince and Smolensky, 1993) analysis for these cases based on syllable weight and double voiced obstruent constructions. Section 6 contains a list of constraints that I use throughout this paper for my OT analyses.

2 Phonemic inventory of Japanese

The native phonemic inventory of Japanese consists of singleton consonants (like /p/, /t/, /k/) as well as geminate consonants (like /pp/, /tt/, /kk/). As shown by the minimal pair data in 1, they are contrastive, and thus separate phonemes. (Kubozono, 2009, p. 2)

(1)   a   /kita/ 'came' vs. /kitta/ 'cut (past)'

        b   /saki/ 'point' vs. /sakki/ 'a short time ago'

However, native Japanese phonotactics disallow the gemination of certain consonants. In particular, the gemination of voiced obstruents (/b/, /d/, /g/, /z/, etc.) is prohibited. The data in sections 2.1 and 2.2 present some evidence for the prohibition of voiced geminates in the native phonology of Japanese.

2.1 Emphatic gemination

In Japanese, gemination sometimes occurs in native words to express emphasis. As shown in 2, /d/ becomes devoiced /tt/ when geminated to avoid */dd/. (Kubozono, 2009, p. 3)

(2)   /tada/ 'only' → [tatta]   *[tadda]

2.2 Gemination from *-ri* suffixation

The suffix *-ri* when attached to mimetic roots causes gemination of the root-final consonant. This gemination is blocked by nasal insertion when the root-final consonant is a voiced obstruent, as in 3b. (Kawahara, 2015, p. 1)

(3)   a   /uka + ri/ → [ukkari] 'absent-mindedly'

        b   /uza + ri/ → [un̩zari] 'sick of something'   *[uzzari]

3 Lexical strata of Japanese

Despite the prohibition of voiced geminates in the native Japanese lexicon, they do appear in the loanword inventory of Japanese, referred to by Ito and Mester (1999) as the FOR-

EIGN strata of the Japanese lexicon. The different lexical strata of the Japanese lexicon as presented by Ito and Mester (1999) are illustrated in 4. Of these four lexical strata, the NATIVE stratum is the only closed set: No new lexical items can be added, and the ones that already exist obey the prohibition of voiced geminates. This constraint is seemingly obeyed for the NATIVE stratum and below, excluding only *Unassimilated foreign*, for which voiced geminates are allowed.

(4) NATIVE

SINO-JAPANESE

FOREIGN

- *Assimilated* ← voiced geminates disallowed

- *Unassimilated* ← voiced geminates allowed

Generally speaking, the separation of the FOREIGN strata into both *Assimilated* and *Unassimilated* variants is to categorize foreign loans into two separate groups: One group for which voiced geminates can be pronounced as devoiced (*Assimilated*), and another for which voiced geminates are mandatory (*Unassimilated*), assuming in both cases that the context would in fact license voiced gemination. From this, it follows that voiced geminates in Japanese are more marked than voiceless geminates, so I use the markedness constraint *VOIGEM.

Example 5 (Kawahara, 2015, p. 3) shows some Japanese loanwords that contain voiced geminates.

(5)   a   'red' → /reddo/

       b   'dog' → /doggu/ (→ /dokku/)

There exist some loanwords for which the voiced geminate consonant can optionally be pronounced as devoiced, as in 5b, and in such cases, this is the form that ultimately assimilates into the lexicon (which likely means that this is a form in the process of lexical assimilation, and thus making the transition from *Unassimilated* → *Assimilated*). This

tendency toward voiceless geminates is evidence for a highly ranked *VoiGem constraint, and the tableau in 6 shows this constraint being used to account for the winning candidate, [tatta], from 2. Faith(Voice) is outranked by *VoiGem, which allows for the output of the unfaithful form with the voiceless geminate.

(6)

/tada/ (emph)	*VoiGem	Faith(Voice)
a. tadda	*!	
☞ b. tatta		*

4   Gemination licensing & motivation

The environment in which consonant gemination in English borrowings occurs is word-finally, when the preceding vowel is lax (or short), as in 7 (Kawahara, 2015, p. 2). However, in the case of a long vowel in the phonemicized form, as in 8, gemination does not occur (Kubozono, 2009, p. 11).

(7)   Gemination

    a   'cat' → [kjatto]

    b   'pack' → [pakku]

    c   'red' → [reddo]

    d   big → [biggu]

(8)   No gemination

    a   'meat' → [miito]   *[miitto]

    b   'peak' → [piiku]   *[piikku]

    c   park → [paaku]   *[paakku]

In order to account for gemination licensing in loanwords, I give two accounts for its motivation in different contexts. The first is a discussion of syllable weight in 4.1, and the second a brief discussion of Lyman's Law and the avoidance of double voiced obstruent constructions in 4.2.

4

4.1 Syllable weight

Japanese words display a tendency to end in Heavy-Heavy (hereafter HH) or Heavy-Light (hereafter HL) syllable sequences, with Light-Light sequences being a close third. The data in 9 shows some examples of baby speech and illustrates this tendency, as LL input words are changed in the output forms to HL words. The tendency is further supported by the process of loanword truncation, as in the examples in 10, where HL truncations are acceptable, but LH are converted to LL, importantly illustrating the intolerance of LH. For these reasons, I use the constraint PROSODICFORM to account for this tendency. (Kubozono, 2009, p. 5)

(9) Baby words (LL → HL)

    a ba.ba → baa.ba 'grandmother'

    b ku.tu → kuk.ku 'shoes'

    c ne.ru → nen.ne 'to sleep'

(10) Loanword truncation (HL)

    a roo.tee.sjon → roo.te 'rotation'

    b pan.hu.ret.to → pan.hu 'pamphlet'

    Loanword truncation (LH → LL)

    c ro.kee.sjon → ro.ke 'location'   *ro.kee

Japanese disallows closed syllables (CVC) for the most part (Shinohara, 1996, p. 7), but they still need to be expressed in the output when the source borrowed word contains them, because MAX(C) is ranked high in my analysis, meaning consonants should not be deleted in the output forms.

The flowchart in 11 illustrates the process for the phonemicization of the English word 'pack' into a Japanese loanword, and its accompanying tableau is in 12. (*Note*: The first half of a geminate obstruent can be represented as /Q/, and in terms of weight, it counts as one mora (μ). Thus, 7b can also be written as [paQ.ku].)

(11)   i   'pack' /pak/ is a *CVC syllable

$$\Downarrow$$

    ii   /u/ is epenthesized to break this into CV.CV → /pa.ku/

$$\Downarrow$$

    iii   /pa.ku/ is *LL word-finally

$$\Downarrow$$

    iv   moraic obstruent /Q/ is added to syllable 1 to increase moraicity → /pa + Q/

$$\Downarrow$$

    v   [paQ.ku] is HL word-finally; PROSODICFORM is satisfied

(12)

/pak/ 'pack'	MAX(C)	PROSODICFORM	NOCODA	DEP(V)	DEP(μ)
a. pak			*!		
b. paQk			*!		*
☞ c. paQ.ku				*	**
d. pa.Qku		*!		*	**
e. pa.ku		*!		*	*
f. pa	*!				

Candidates (a) and (b) are ruled out by violations of NOCODA, which are equal in weight to the violations of both PROSODICFORM and MAX(C). Candidates (d) and (e) incur violations of PROSODICFORM, with (d) being LH word-finally, and (e) being LL. Candidate (f) violates MAX(C), since its coda-position /k/ was deleted. The low-ranking constraints DEP(V) and DEP(μ) are violable, as the winning candidate (c) incurs a violation of DEP(V) via /u/-epenthesis, and two violations of DEP(μ): one via /u/-epenthesis, and one via μ-epenthesis.

A note on /Q/: I am making the assumption that the moraic obstruent /Q/ does not count as a coda on its own, but does however increase moraicity of the containing syllable, and thus contributes to the satisfaction of PROSODICFORM. The evidence for this is possibly

the non-realization of /Q/ as a phonetic segment, and instead it is a theoretical unit that anchors to the following consonant, and adds a mora (μ). It is for this reason that candidate (b) in the tableau in 12 incurs only one violation of NoCoda, but two violations of Dep(μ).

Now recall example (b) from 8 above:

(8) b   'pea**k**' $\rightarrow$ [pii.**k**u]   *[pii**Q.k**u]

The question now is Why is gemination not licensed in this case? According to my analysis with syllable weight and the ProsodicForm constraint from earlier, the word-final /k/ in 8b should be allowed to geminate, resulting in [piiQ.ku]. This is not the attested output, however. Thus, it either seems insufficient to say that word-final HH / HL forms are desired, or "Heavy" syllables need to be more strictly defined. If [piiQ] were not considered a Heavy syllable, the gemination would be blocked as we see in the desired output. For this reason, I use the constraint *Superheavy to account for the prohibiton of Superheavy syllables containing three moras. In other words, long vowels count as two moras making Japanese a quantity sensitive language in the context of syllable weight.

This constraint is put to use in the tableau in 13 to account for the winning candidate, [pii.ku]. Candidates (a) and (b) are ruled out via NoCoda violations. Candidate (e) is ruled out via a Max(C) violation, since its coda /k/ was deleted. Candidates (c) and (d) are tied at this point, since neither violate ProsodicForm, but (c) is ultimately ruled out via a *Superheavy violation from its trimoraic first syllable, /piiQ/.

(13)

/piik/ 'peak'	Max(C)	ProsForm	NoCoda	*Superheavy	Dep(V)	Dep(μ)
a. piik			*!			
b. piiQk			*!			*
c. piiQ.ku				*!	*	**
☞ d. pii.ku					*	*
e. pii	*!					

4.2 Lyman's Law and the D2 constraint

According to Rice (2006), there is a constraint present in Japanese phonology (among other languages) known as Lyman's Law which deals with the well-formedness of morphemes. Specifically, this constraint says that morphemes cannot contain two voiced obstruents, though there are some exceptions. Lyman's Law can be observed operating throughout Japanese phonology by simply looking at the combinations of voiced + voiceless segments possible in native words. This is illustrated in 14 (Rice, 2006, p. 14), where the interactions between [d], [t], [b], and [f] [1] are shown. Crucially, [+voi], [+voi] combinations within a morpheme are not found in this example.

(14)   a   /futa/ 'lid' ([-voi, [-voi])

        b   /fuda/ 'sign' ([-voi], [+voi])

        c   /buta/ 'pig' ([+voi], [-voi])

        d   */buda/ ([+voi], [+voi])

Now recall example (b) from 5 earlier:

(5) b 'dog' $\rightarrow$ /doggu/ ($\rightarrow$ /dokku/)

Previously, I had accounted for this optionality in geminate devoicing by positing that voiced geminates are inherently more marked than voiceless geminates, and proposing the use of the constraint *VoiGem. Now with the introduction of Lyman's Law, I can provide an alternative account for the tendency of speakers to devoice the geminate from 5b using the tableau in 15. The markedness constraint *D2$_m$ (Ito and Mester, 2003, p. 37) says that there should not be two voiced obstruents within a morpheme. This successfully rules out candidate (a) [doggu] (which contains two voiced obstruents within a morpheme), and accounts for the winner, [dokku].

(15)

/doggu/ 'dog'	*D2$_m$	FAITH(VOICE)
a. doggu	*!	
☞ b. dokku		*

---

[1][f] is one of the allophones that functions as the voiceless counterpart of /b/ in Japanese

The tableau in 16 shows an example (5a) for which this analysis successfully predicts the voiced geminate winner, since there is no *D2$_m$ violation and thus no motivation for devoicing of the geminate.

(16)

/reddo/ 'red'	*D2$_m$	FAITH(VOICE)
☞ a. reddo		
b. retto		*!

5 Conclusion

Geminates in Japanese loanwords provide a source of data that allows for inspection of issues that arise mainly in the non-native phonotactics of Japanese, and also allows us to analyze these phenomena alongside the native phonotactics. Section 2 introduced the prohibition of voiced geminates in the native phonology of Japanese. Then, in section 3, the discussion of the different lexical strata of Japanese showed that voiced geminates are seemingly more marked than voiceless ones. Section 4 presented an analysis of syllable weight and Lyman's Law in Japanese, and how they inspire gemination and geminate devoicing respectively. Syllable weight plays an important role in the licensing of gemination in loanwords, and it seems that Japanese is very sensitive to syllable weight. Further research about borrowings from languages other than English into Japanese might allow for alternative analyses of how syllable weight is encoded in both the source words and in the Japanese phonemicizations.

6 Constraints

*VOIGEM: Voiced obstruent geminates are prohibited

FAITH(VOICE): Don't change voicing values from the input

PROSODICFORM: Words with two or more syllables must end in HH or HL sequences

MAX(C): Don't delete consonants in output forms

NOCODA: Output forms must not have codas

DEP(V): Don't epenthesize vowels in output forms

Dep(μ): Don't increase moraicity in output forms

*Superheavy: Trimoraic syllables are banned

*D2$_m$: Output forms must not have two voiced obstruents within a morpheme

References

Ito, J. and Mester, A. (1999). The phonological lexicon. *A Handbook of Japanese Linguistics*.

Ito, J. and Mester, A. (2003). *Japanese Morphophonemics*. MIT Press, Cambridge, Massachusetts.

Kawahara, S. (2015). Geminate devoicing in japanese loanwords: Theoretical and experimental investigations. *Language and Linguistics Compass*, 9.

Kubozono, H. (2009). Consonant gemination in japanese loanword phonology. *Current Issues in Unity and Diversity of Languages, The Linguistic Society of Korea*, pages 953–973.

Prince, A. and Smolensky, P. (1993). Optimality theory: Constraint interaction in generative grammar. *Rutgers University Center for Cognitive Science*, 2.

Rice, K. (2006). On the patterning of voiced stops in loanwords in japanese. *Toronto Working Papers in Linguistics*, 26:11–22.

Shinohara, S. (1996). The roles of the syllable and mora in japanese: adaptation of french words. *Cahiers de linguistique - Asie orientale*, 25:87–112.