

The core–periphery model of the lexicon and phonotactic constraints on loanwords in Japanese

Clifford Crawford
Department of Linguistics
Cornell University
Ithaca, NY 14850
<cjc26@cornell.edu>

Introduction

Itō and Mester (1995, 1999) have proposed the *core–periphery* model of the lexicon, in which the lexicon of a given language is organized in terms of overlapping domains of markedness constraints. In this model, the native vocabulary exists at the “core” of the lexicon, satisfying all possible markedness constraints, while loanwords enter the lexicon in the outer “periphery” where fewer constraints are obeyed, and gradually move inward toward the core as they become nativized over time. While Itō and Mester have developed the core–periphery model primarily using data from Japanese, this model has become the dominant way of explaining the behavior of loanword adaptation in many other languages, for example in analyses of English loanwords in Québécois French (Paradis and Lebel 1997), Korean (Cho 1999), and American Sign Language (Brentari and Padden 2001). However, I will argue that the core–periphery model, when applied to Japanese, is too simplistic, for the following two reasons: (1) there is no way to consistently rank all of the markedness constraints governing the lexical strata of Japanese; (2) the model cannot explain the existence of constraints on recent loanwords (in particular, constraints on word length and on the distribution of palatalized consonants) which do not govern the other strata.

Lexical strata in Japanese

The existence of distinct strata in the lexicon of Japanese has long been recognized (Martin 1952, McCawley 1968, Itō and Mester 1995), and is somewhat analogous to the phenomenon of lexical strata in other languages, such as level 1 and level 2 morphology in English (Borowsky 1986). Japanese is usually described as having four distinct strata. The native vocabulary which can be traced back to Old Japanese is known as *Yamato* (from a Japanese word meaning ‘native’). Yamato items are extremely common in natural speech and in various genres of written text (Shibatani 1999, pp. 142–3). The large set of loanwords which were borrowed from Middle Chinese in the 7th–13th c. is known as *Sino-Japanese*. These words comprise about 50% of the total lexicon, and are somewhat analogous in status to the Latinate vocabulary of English (pp. 145–6). Loanwords which were borrowed from the 15th c. onward comprise the *Foreign* stratum. Almost all Foreign words come from English, French, German, and other European languages, with English dominating. Finally, there is a semi-productive set of sound-symbolic words known as the *Mimetic* stratum (Hamano 1998).

Although these strata are defined mostly by their etymology, they also obey distinct sets of phonotactic constraints which interact with various morphological processes. For example, a constraint holding in the Yamato stratum is *NT, which disallows voiceless stops following nasals. *NT can be seen to operate when the gerund suffix /--te/ or the past tense suffix /--ta/ is attached to verb bases ending in a nasal, causing the /t/ in both suffixes to be voiced (Figure 1). This constraint does not hold for the Sino-Japanese stratum, however, and there are numerous examples of post-nasal voiceless stops in Sino-Japanese words, for example *benkjo*: ‘study’.

Both Yamato and Sino-Japanese also obey a constraint *P which prevents an underlying /p/ from surfacing when it is not a geminate or not following a nasal. This rather unusual constraint came about as a result of a sound change which affected word-initial and intervocalic *p, leniting it to *φ. Word-initial *φ then became /h/, except before /u/, where it remained /φ/. Intervocalic *φ, however, became *w (again,

	root	negative --(a)nai	gerund --te	past --ta
‘see’	mi--	minai	mitte	mita
‘buy’	kaw--	kawanai	katte	katta
‘die’	jin--	jinanai	jinde	finda
‘read’	jom--	jomanai	jonde	jonda

Figure 1: Past tense alternations in the Yamato stratum

except before /ɰ/), which was later deleted, except before /a/ (Shibatani 1999, p. 167). This diachronic relationship between /h/ and /p/ is reflected in several morphophonemic alternations between /h/ and /p/ (or /b/). For example, some number words like itʃi ‘one’ and dʒiu: ‘ten’ cause gemination when prefixed to a counter word, as in issatsuu < itʃi+satsuu ‘one volume’, dʒuussatsuu < dʒiu:+satsuu ‘ten volumes’. But when these words are prefixed to the counter word hon, a geminate /p/ shows up instead: ippoɴ, dʒuuppon. In this case, the underlying /p/ only surfaces when it is geminated, and not when it would occur word-initially, because it would violate *P.¹

Neither *NT nor *P govern the Foreign stratum, though. There are numerous examples of voiceless stops occurring after nasals, both across morpheme boundaries (as in dauntaimuu < daun+taimuu ‘down time’) and word-internally (paɴku ‘punk’; inta:netto ‘internet’). And /p/ is unrestricted in loanwords, occurring not only in geminates or nasal clusters, but also word-initially (purosesuu ‘process’; peɴ+paruu ‘pen pal’) and intervocally (apato ‘apartment’, taipuu ‘type’). *P is also violated in the Mimetic stratum; for example, /p/ occurs frequently as the initial consonant in CVCV Mimetic roots (Hamano 1998, p. 41) in words like poroporo ‘in large drops’ or pikapika ‘glitter; sparkle’. *NT is obeyed in the Mimetic stratum, but only within roots, not across morpheme boundaries, as shown by forms like pimpin < reduplicated pin ‘lively’ or dosumto < dosuɴ+to ‘with a thump’.

The core-periphery model of the lexicon

Itō and Mester (1999) claim that these different morphophonemic properties of the different strata show that the Japanese lexicon has a *core-periphery* structure. In this model, Yamato forms the core of the lexicon, because it is governed by the most constraints, while Foreign forms the outermost periphery of the lexicon because it obeys the fewest constraints (basically, only constraints on syllable structure which are obeyed in all strata). Sino-Japanese falls in between these two layers. In particular, in Yamato *P, *NT, and CODACOND (coda consonants may not have place features) all hold, in Sino-Japanese only *P and CODACOND hold, and in Foreign only CODACOND holds (Figure 2). This suggests that the strata are arranged in a subset relationship, with the set of possible Sino-Japanese words forming a subset of possible Foreign words, and the set of possible Yamato words forming a subset of possible Sino-Japanese words. This mirrors the etymological history of the different strata—Yamato items having always been a part of the language, Sino-Japanese items having been borrowed relatively long ago, before the 14th c., and Foreign items having been borrowed relatively recently. Itō and Mester (1995) argue that this core-periphery structure reflects the different degrees of nativization that the different strata have undergone, Foreign items having been nativized less than Sino-Japanese items because they were borrowed later. They also claim that Foreign items do not obey any constraints which are not also obeyed in the other strata, and argue because of this that Foreign should not be considered a real stratum at all (p. 824).

¹McCawley (1968) proposes to eliminate the /h/ phoneme from the Yamato and Sino-Japanese strata entirely, deriving all instances of /h/ from an underlying /p/. This proposal is not widely accepted, however, because it seems rather counterintuitive to native speakers, and because the change from /p/ to /h/ is highly unusual, from a phonetic viewpoint (Shibatani 1999, p. 167). Also, both /h/ and /p/ can occur word-initially and intervocally in loans, and so cannot be in an allophonic relationship in the Foreign stratum.

	Yamato	Sino-Japanese	Mimetic	Foreign
CODACOND	✓	✓	✓	✓
*P	✓	✓	violable	violable
*NT	✓	violable	✓	violable
ROOT $\leq 2\mu$	violable	✓	violable	violable
ROOT = 1ϕ	violable	violable	✓	violable

Figure 2: Pattern of violations of various phonotactic constraints in Japanese

One problem with the core-periphery model, as it is applied to Japanese, is that the facts are less clear when we take the Mimetic stratum into account. Recall that Mimetic words can violate *P, but not *NT. Although the set of possible Sino-Japanese and Mimetic items both form a subset of possible Foreign words, it is not possible to place Sino-Japanese and Mimetic in a subset relationship with each other, because there are some Sino-Japanese items which violate *NT and hence cannot be possible Mimetic words, and likewise there are some Mimetic items which violate *P and cannot be possible Sino-Japanese words.

Even if the relationship between Sino-Japanese and Mimetic is unclear, we could maintain that Yamato still serves in some sense as the core of the lexicon, because it is more constrained in terms of possible phonotactic patterns than any of the other strata, at least with respect to the constraints we've already seen. Yet there do exist constraints which govern only Sino-Japanese or Mimetic as well. Sino-Japanese roots can only be 1 or 2 morae long (Tateishi 1990), while Mimetic words are always formed from bimoraic (or trimoraic) roots which form a prosodic foot (Poser 1990). (These are indicated by $ROOT \leq 2\mu$ and $ROOT = 1\phi$, respectively, in Figure 2.) The existence of these constraints is completely unaccounted for in the core-periphery model. Yamato, being at the core of the lexicon, should be governed by all of the constraints which govern the other strata, yet these length restrictions on Sino-Japanese and Mimetic do not apply to Yamato, as shown by the existence of Yamato words that are only one mora long, like *ki* 'tree; wood' and *me* 'eye', as well as the large number of Yamato roots that are more than two morae long.

Thus, the core-periphery model is not a good model of the relationships between the Yamato, Sino-Japanese, and Mimetic strata, because these strata cannot be placed in a subset relationship with respect to the phonotactic constraints they obey. The question now is whether the predictions of the core-periphery model concerning the Foreign stratum are also true: namely, that it is not a real stratum at all, but simply a heterogeneous collection of loanwords in various stages of nativization, with no systematic phonotactic patterns. To do this, we need to look for possible phonotactic constraints which govern the Foreign stratum, but which are freely violated in the other strata. If there do exist such constraints, then this is another point of evidence that the Japanese lexicon is not organized in an overall core-periphery structure.

Methods

JMDICT (EDRDG 2003), an electronic Japanese-English dictionary with over 90,000 entries, was used to conduct the analysis. Each entry was classified according to stratum using a computer program. The Yamato, Sino-Japanese, and Foreign strata are easily distinguished by the scripts normally used for each: *katakana* for Foreign, *kanji* for Sino-Japanese, and *hiragana* and *kanji* for Yamato. Moreover, each *kanji* character has two sets of readings, the Sino-Japanese reading, or *on-yomi*, and the Yamato reading, or *kun-yomi*. (For example, the character 山 is used both for the Sino-Japanese root *san* and the Yamato word *jama*, both meaning 'mountain'.) A word written in *kanji* can thus be classified as Sino-Japanese or Yamato based on whether the *on-yomi* or *kun-yomi* readings of the characters are used.

Unfortunately, there is no way to distinguish the Mimetic items using orthographic criteria, because in JMDICT they are written using either *hiragana* or *katakana*. However, Mimetic words are usually formed from bimoraic roots, which are either reduplicated or occur with a suffix like /-ri/, /-tsukuu/, or /-tosuruu/ (Hamano 1998), and so all words with these properties were classified as Mimetic. There were

Yamato	16,710	(16.9%)
Sino-Japanese	46,132	(46.8%)
Mimetic	736	(0.7%)
Foreign	14,500	(14.7%)
Hybrid	20,588	(20.9%)
Total	98,666	

Figure 3: Lexical strata derived from JMDICT

also many entries which were compounds formed from items in different strata. These were categorized as Hybrid and were not considered in the analysis below.

Figure 3 summarizes the results of the stratum–assignment algorithm applied to the entries in JMDICT. The proportions of Yamato, Sino-Japanese, and Foreign words in the dictionary are comparable to the statistics cited in Shibatani (1999, pp. 142–3), except that there are too few Yamato words (compared to about 30–40% in the texts Shibatani cites), and too many Hybrid and Foreign words (about 10% for each in Shibatani). Presumably this is because the dictionary lists many compounds and foreign borrowings which are not commonly found in real texts, and so this skews the ratios towards these types of items. Also, the number of Mimetic items found is extremely small, because such terms are usually not listed in Japanese dictionaries (Hamano 1998) and because of the difficulty identifying them as noted above. Since there was not a large enough sample of Mimetic words, I used only the words classified as either Yamato, Sino-Japanese, or Foreign in the analysis that follows.

After running the stratum–classification program on the entries in JMDICT, I then looked for evidence of constraints unique to the Foreign stratum by looking for phonotactic patterns which are rare in the Foreign stratum but common in the Yamato and Sino-Japanese strata. I then investigated each candidate pattern to see if it was regularly avoided in the process of loanword adaptation. This is important because the rarity of a pattern in the dictionary is not sufficient proof for the existence of a constraint against it. For instance, voiced obstruents following a nasal turn out to be almost as rare as voiceless obstruents in the same environment in the set of Yamato words in JMDICT, yet /nd/ sequences regularly occur in the past tense and gerund forms of verb roots ending in a nasal (Figure 1), making it unlikely that there actually is a constraint ruling out such sequences.² Two candidate constraints were identified: a minimal–length constraint, and a constraint on the distribution of palatalized consonants.

Minimal length constraints on loanwords

First of all, there seems to be a minimal length constraint on possible Foreign words, similar to the length constraints that govern the Sino-Japanese and Mimetic strata. Both Lovins (1975) and Tsuchida (1995) note that the tense–lax distinction in English vowels is usually represented by length in the corresponding borrowings in Japanese: tense vowels are borrowed as long (= 2μ) vowels, while lax vowels are borrowed as short (= 1μ) vowels. In addition, because Japanese does not allow coda consonants in general (with the exception of moraic nasals, and the first segment of a geminate), such consonants in the source word will have epenthetic vowels (usually /u/) inserted after them.³ This results in every loanword having at least two morae. Some examples of loanwords derived from monosyllabic English words are given in Figure 4. In every case, the loanword is at least two morae long. For the source words with tense vowels this follows directly from the fact that tense vowels are borrowed into Japanese as long vowels. For the source words

²This is a disadvantage of using a dictionary for investigating phonotactic constraints: dictionaries do not normally list all of the outcomes of regular morphological processes, such as the various inflected forms of verbs. I would expect that the frequency of /nd/ sequences in a sample of Yamato items taken from a corpus of written text, such as a collection of newspaper articles, would be much higher.

³Under certain conditions, the consonant will be geminated as well. See Lovins (1975) and Tsuchida (1995) for details.

Open σ , tense V	Closed σ , lax V	Closed σ , tense V
'key' > ki:	'lip' > rippu	'cheap' > tʃi:pʊ
'pay' > pe:	'pet' > petto	'cape' > ke:pʊ
'show' > ʃo:	'loss' > rosu	'rope' > ro:pʊ
'you' > yʊ:	'put' > putto	'hoop' > ɸu:pʊ

Figure 4: Adaptation patterns of English monosyllables

2 morae

kʲara	< kʲarakʲuta:	'character'
tʃoko	< tʃokore:to	'chocolate'

3 morae

arumi	< aruminʲu:mʊ	'aluminum'
terebi	< terebidʒoN	'television'

4 morae

pasokon	< pa:sonaru+kompʲu:tʲa:	'personal computer'
apato	< apato:mento	'apartment'

Figure 5: Examples of truncated loanwords

with lax vowels, which are always closed syllables in English,⁴ the second mora comes from the epenthetic vowel inserted after the coda consonant.

A similar pattern arises in truncated versions of loanwords (Figure 5). Usually these are formed from the first two morae of the loanword. Although there are also some examples that are three or four morae long, there are never any that are only one mora long (Itō 1990, Labrune 2002). Both of these facts suggest that there is some kind of minimal length constraint operating on words in the Foreign stratum:

MINWORD: A word must have at least two morae.⁵

There are only three loanwords listed in JMDICT which violate MINWORD. One of them, *za* 'the', although listed as a separate word, actually only occurs as a bound morpheme in titles and such, so it is not a true exception. The other two exceptions, *tī* 'tea' and *de* 'day', have two-mora variants, *tī:* and *de:*, which are more common (John Whitman, p.c.). Also, Lovins (1975) notes that one-mora loans are short-lived and quickly replaced by two-mora variants created by lengthening the vowel, which is probably what is happening to *tī* and *de*. While one-mora words rarely, if ever, occur in the Foreign stratum, they do occur much more frequently in Sino-Japanese and Yamato, and include such common words as *ki* 'tree; wood', *te* 'hand', *çi* 'sun; sunshine; day', and *me* 'eye'. Thus MINWORD is an example of a constraint which applies in the Foreign stratum, but not in Yamato or Sino-Japanese.

Palatalized consonants in the Foreign stratum

Next I turn to the distribution of palatalized consonants. First note that palatalized /r/ occurs only before /u/ in loanwords (Figure 6). The only exception listed in JMDICT is *kʲasuterʲa* 'Castilian'. Also, every

⁴Except for function words.

⁵For ease of exposition I have formulated this as a single constraint, but Kager (1999) suggests that minimal length effects actually arise from the interaction of two constraints, FTBIN (feet are binary in terms of syllables or morae) and GRWD=PRWD (a grammatical word is a prosodic word).

ir ^j u:ɖʒoN	‘illusion’
ebar ^j u:e:fɔN	‘evaluation’
k ^j asuter ^j a	‘Castilian’
suk ^j u:r ^j u:	‘screw’
so:r ^j u:fɔN	‘solution (esp. to a problem)’
to:r ^j uφu	‘truffle’ (from French <i>truffe</i>)
bar ^j u:	‘value’
φe:r ^j ua	‘failure’
pu:r ^j u:do	‘prelude’
po:r ^j u:fɔN	‘pollution’
bo:r ^j u:m ^j u	‘volume’
r ^j u:ɖʒu	‘sledge’ (from French <i>luge</i>)
rezo:r ^j u:fɔN	‘resolution’

Figure 6: Examples of Foreign items with /r^j/

aru:man ^j akk ^j u	‘armagnac’
esu:pan ^j o:r ^j u:so:s ^j u	‘sauce Espagnole’
eru:mi:n ^j o	‘El Nino’
gin ^j o:r ^j u	‘guignol’
k ^j asuter ^j a	‘Castilian’
ku:ro:man ^j oN	‘Cro-Magnon’
ko:n ^j akk ^j u	‘cognac’
fin ^j o:nsu:ta:ru	‘chignon style (hairdo)’
m ^j amma:	‘Myanmar’
ra:za:m ^j a	‘lasagna’
ro:fin ^j o:r ^j u	‘Rossignol’

Figure 7: Loanwords with palatalized consonants occurring before /o/ or /a/ (does not include velar stops)

one of these /r^ju/ sequences corresponds to /l(j)u/ or /r(j)u/ in the source word,⁶ although /ru/ and /lu/ can also be borrowed as just /ru/, as in bu:ru: ‘blue’ and ru:to ‘root; route’. But, /ljə/⁷ is borrowed as /ri(j)V/, not /r^jV/, as in mirioN ‘million’ and biri:ja:do ‘billiards’. Palatalized fricatives and nasals in the Foreign stratum show this same distribution, as well as labial and alveolar stops. (Velar stops, however, can be palatalized before /a/ as well as /u/, a point I will return to shortly.) With only a few exceptions listed in JMDICT (Figure 7), they occur only before /u/.⁸ To explain this distribution, I propose that the following two constraints are operating in the Foreign stratum:

***PALMID:** Palatalized consonants cannot occur before mid vowels.

***PALLOW:** Palatalized consonants cannot occur before low vowels.

⁶One might argue that, since there are dialectal differences in whether the /l/ in words like ‘revolution’ is palatalized or not, then Japanese must be borrowing these words only from the palatalizing dialects of English. However, this fails to explain the /r^j/ in suk^ju:r^ju: ‘screw’ and related forms, since as far as I know there is no dialect of English in which ‘screw’ is pronounced with a palatalized /r/.

⁷There don’t seem to be any good examples of /rjə/ in English; ‘carrier’ and ‘warrior’ are marginal cases where /ri/ surfaces as /rj/ in rapid speech, and ‘carrier’ is listed in JMDICT as k^jaria, with /ri/ instead of /rj/.

⁸Nearly all of these exceptions come from /n^j/ spelled <gn> in words borrowed from French, suggesting that there might be some sort of spelling-based adaptation going on here. /n^j/ when it is *not* spelled as <gn> in the source word is consistently borrowed as /n^j/ before /u/ and /ni/ otherwise, as in words like k^jani:ɔN ‘canyon’ and kenia ‘Kenya’.

/bar ^j u:/	MAX-IO[FRONT]	DEP-IO	*PALMID	CONTIG-IO
☞ [bar ^j u:]				
[bariu:]				*!
[baru:]	*!			

/mir ^j oN/	MAX-IO[FRONT]	DEP-IO	*PALMID	CONTIG-IO
[mir ^j oN]			*!	
☞ [mirioN]				*
[miron]	*!			
[mir ^j uoN]		*!		*

Figure 8: ‘value’ > bar^ju:, *bar(i)u:; ‘million’ > mirioN, *mir^joN

ka:ɸu	‘calf’	suk ^j aN	‘scan’
kadzuaru	‘casual’	k ^j assuru	‘castle’
katapira:	‘caterpillar’	k ^j atto	‘cat’
kategori	‘category’	g ^j appu	‘gap’
kamera	‘camera’	k ^j andi:	‘candy’
karenda	‘calendar’	g ^j aroppu	‘gallop’

Figure 9: Examples of loans from source words with a velar stop before /æ/

These two constraints are similar to another constraint on palatalized consonants which governs the entire Japanese lexicon:

***PALFRONT:** Palatalized consonants cannot occur before front vowels.⁹

Suppose *PALMID, MAX-IO[FRONT] (preserve palatalization of consonants, assuming palatalization is represented by a [+front] feature), and DEP-IO (no epenthesis) are all ranked higher than a constraint CONTIG-IO requiring contiguous segments in the input to correspond to contiguous segments in the output (Kager 1999, p. 250). Then this ranking would predict that palatalized consonants can surface before /u/, but not before /o/,¹⁰ and in the latter case, the violation of *PALMID will be repaired not through depalatalization or epenthesis, but by creating two segments in the output corresponding to and sharing the features of the palatalized consonant in the input, so that e.g. /r^j/ in the input corresponds to /ri/ in the output (Figure 8).¹¹

Ranking *PALLOW above CONTIG-IO would also work to explain why palatalized consonants also can’t appear before /a/. However, there is a major exception to *PALLOW, namely that English words with a velar stop followed by /æ/ are often borrowed with palatalized velars as well as with plain velars (Figure 9). In fact, /k/ or /g/ before /æ/ in a source word is just as likely to be palatalized in the resulting loanword as not; JMDICT lists approximately 100 loanwords having plain velars in word-initial position, and about the same number having palatalized velars. Also, whether velars in this environment will be palatalized or not seems to be rather unpredictable, given the existence of pairs like k^jatto ‘cat’, borrowed with a palatalized /k/, and katapira: ‘caterpillar’, with a plain /k/, where in both cases the /k/ occurs in the same phonological environment in the source word (in a stressed syllable before /a/).

⁹Unlike most analyses of Japanese phonology, I am not considering the affricates /tʃ/ and /dʒ/ to be the palatalized versions of /t/ and /d/, because in the Foreign stratum there do exist examples of /tʃ/ and /dʒ/, for instance d^juti: ‘duty’.

¹⁰Or /e/, the other mid vowel in Japanese, but this is already ruled out by the highly ranked *PALFRONT.

¹¹Note that there do exist a few loanwords which show epenthesis in this context, such as ɸijoruudo ‘fjord’.

/k ^j atto/	*PAL[-VEL]LOW	MAX-IO[FRONT]	CONTIG-IO	*PAL[+VEL]LOW
☞ [k ^j atto]				*
[kiatto]			*!	
[katto]		*!		

/ken ^j a/	*PAL[-VEL]LOW	MAX-IO[FRONT]	CONTIG-IO	*PAL[+VEL]LOW
[ken ^j a]	*!			
☞ [kenia]			*	
[kena]		*!		

Figure 10: ‘cat’ > k^jatto, *k(i)atto; ‘Kenya’ > kenia, *ken^ja

There seem to be two things going on here. One is that Japanese speakers, for some reason, often perceive English velar stops before /æ/ as palatalized. This is probably due to the frontness of /æ/, and corresponds to a well-known allophonic variation in English between velars before back vowels, with a relatively back place of articulation (as in the word ‘coo’), and velars before front vowels, with the tongue body making contact further front, near the soft palate (as in ‘key’) (Keating and Lahiri 1993). However, this perception-based account would predict that any front vowel in the source word should cause palatalization of /k/ and /g/; but, in fact only /æ/ does so. This is because, as noted before, there is a highly-ranked constraint *PALFRONT (no palatalized consonants before front vowels) operating in the entire Japanese lexicon, and so sequences like /k^je/ or /k^ji/ would not be allowed to surface. /æ/ is unique among English front vowels in being borrowed into Japanese not as a front vowel, but as a central vowel /a/. Thus /kæ/ > /k^ja/ would not violate *PALFRONT. Also, because /æ/ is the least front of the front vowels in English, it would cause less fronting on a preceding velar than /i/ or /e/ would, and so it is possible that Japanese speakers do not always perceive velars before /æ/ as palatalized. This could explain the apparently random variation in borrowing sequences like /kæ/ as either /k^ja/ or /ka/, with either a palatalized or plain velar.¹²

Yet sequences like /k^ja/ do violate another constraint I proposed, *PALLOW (no palatalized consonants before low vowels). There doesn’t seem to be a good solution to this, other than to split *PALLOW into two constraints, one for velar consonants (*PAL[+VEL]LOW) and one for non-velars (*PAL[-VEL]LOW), with the ranking *PAL[-VEL]LOW, MAX-IO[FRONT] >> CONTIG-IO >> *PAL[+VEL]LOW (Figure 10). The final ranking, then, is:

MAX-IO[FRONT], DEP-IO, *PALFRONT, *PALMID, *PAL[-VEL]LOW >>
CONTIG-IO >> *PAL[+VEL]LOW

But while violations of *PALMID and *PALLOW are normally avoided in loanword adaptations, they are frequently violated in the Sino-Japanese stratum, as shown by the frequent occurrence of palatalized consonants before /o/ and /a/ in words like beŋk^jo: ‘study’. Thus *PALMID and *PALLOW are two more examples of constraints which are unique to the Foreign stratum.

Conclusion

I have shown that the Japanese lexicon does not exhibit an overall core-periphery structure. The Yamato, Sino-Japanese, and Mimetic strata cannot be organized in an subset relationship with each other, because

¹²Trying to explain the adaptation of /kæ/ as a plain velar /ka/ by invoking *PALLOW would not be a good idea, because the repair strategy (depalatalizing the /k/) is different in this case than in other cases of violation of *PALLOW, which usually show /Cj/ > /Ci/ instead (as in kenia ‘Kenya’).

*NT governs Yamato and Mimetic but not Sino-Japanese, while *P governs Yamato and Sino-Japanese but not Mimetic, and because Sino-Japanese and Mimetic each have word or root length constraints which do not hold in Yamato. Also, there exist at least two sets of phonotactic constraints, MINWORD and PAL[MID,LOW], governing the Foreign stratum which do not also govern Yamato or Sino-Japanese, contrary to what would be predicted for a peripheral layer of the lexicon.

While the Japanese lexicon does not have a core–periphery structure, the concept does seem to work well in characterizing the loanword data and the process of nativization in other languages like Korean (Cho 1999) and Québécois French (Paradis and Lebel 1997). Korean is an especially interesting case to compare to Japanese, because both languages have a similar organization of the lexicon into four strata: native words, mimetics, loanwords from Middle Chinese, and recent loanwords. The difference is, in Japanese, loanwords make up over 10% of the total lexicon, while in Korean, loans only make up about 5% (Sohn 1999). In Canadian dialects of French, loans also make up a relatively small proportion of the lexicon, about 1–5% (Brown 2003). This suggests that the core–periphery model is largely correct for describing the behavior of loanwords in languages in which loans constitute a small fraction of the total lexicon. However, when loanwords make up 10% or more of the lexicon, then it seems to be possible for some of the effects of phonotactic constraints governing the source language to be salient enough in the set of loanwords that these constraints are also considered to be governing loanwords by speakers of the borrowing language. (Note that MINWORD and *PAL[MID,LOW] are both related to similar phonotactic constraints in English on word length and on the distribution of /CjV/ sequences.) In this case, then, the set of loanwords can be considered to be forming a new stratum in the lexicon of the borrowing language, as Foreign loanwords are in Japanese.

References

- Borowsky, T. 1986. Topics in the lexical phonology of English. Ph.D. dissertation, University of Massachusetts.
- Brentari, D., and C. A. Padden. 2001. Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins. In Brentari, D., ed. Foreign vocabulary in sign languages: A cross-linguistic investigation of word formation. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc. pp. 87–119.
- Brown, B. 2003. Code–convergent borrowing in Louisiana French. *Journal of Sociolinguistics* 7, pp. 3–23.
- Cho, Y. Y. 2001. Language change and the phonological lexicon of Korean. In Brinton, L. J., ed. *Historical linguistics 1999: Selected papers from the 14th International Conference on Historical Linguistics*. Amsterdam: John Benjamins. pp. 89–104.
- Electronic Dictionary Research and Development Group. 2003. JMDICT (Japanese–English dictionary). Monash University, School of Computer Science and Software Engineering. <http://www.csse.monash.edu.au/~jwb/j_jmdict.html>
- Fukazawa, H., and M. Kitahara. To appear. Ranking paradox in consonant voicing in Japanese. In van de Weijer, J., K. Nanjo, and T. Nishihara, eds. *Voicing in Japanese*. <http://www.yo.rim.or.jp/~mkitahar/Paper/JVoice_draft.pdf>
- Fukazawa, H., M. Kitahara, and M. Ota. 1998. Lexical stratification and ranking invariance in constraint–based grammars. ROA–267, Rutgers Optimality Archive. <<http://roa.rutgers.edu/>>
- Hamano, S. 1998. *The sound–symbolic system of Japanese*. Stanford, Ca.: CSLI Publications.
- Itō, J. 1990. Prosodic minimality in Japanese. In Ziolkowski, M., M. Noske, and K. Deaton, eds. *The syllable in phonetics and phonology*. CLS 26, vol. 2. Chicago: Chicago Linguistic Society. pp. 213–239.
- Itō, J., and R. A. Mester. 1995. Japanese phonology. In Goldsmith, J. A., ed. *The handbook of phonological theory*. Oxford: Blackwell Publishers Ltd.
- . 1999. The phonological lexicon. ROA–256, Rutgers Optimality Archive.
- Kager, R. 1999. *Optimality theory*. Cambridge University Press.

- Kawahara, S., K. Nishimura, and H. Ono. To appear. Unveiling the unmarkedness of Sino–Japanese. In McClure, W., ed. *Japanese/Korean Linguistics 12*. Stanford, Ca.: CSLI Publications. <<http://www.ags.uci.edu/~hajime/JKProcFin.pdf>>
- Keating, P., and A. Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50, pp. 73–101.
- Labrune, L. 2002. The prosodic structure of simple abbreviated loanwords in Japanese: a constraint-based account. ROA–532, Rutgers Optimality Archive.
- Lovins, J. 1975. Loanwords and the phonological structure of Japanese. Indiana University Linguistics Club.
- Martin, S. E. 1952. Morphophonemics of standard colloquial Japanese. Ph.D. thesis, Yale University.
- McCawley, J. D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.
- Ono, H. 2002. Sino–Japanese and a way to shape up the stems. In *Proceedings of the Southwest Workshop on Optimality Theory*. Department of Linguistics, University of Texas, Austin. <<http://www.ags.uci.edu/~hajime/papers/Ono2002SWOT.pdf>>
- Paradis, C., and É. Lebel. 1997. Centre et périphérie: les segments anglais en français québécois. In Aguer, J. and Y. Rose, eds. *Explorations du lexique*. Québec: International Center for Research on Language Planning. pp. 229–242.
- Poser, W. 1990. Evidence for foot structure in Japanese. *Language* 66, pp. 78–105.
- Rice, K. 1997. Japanese NC clusters and the redundancy of postnasal voicing. *Linguistic Inquiry* 28, pp. 541–551.
- Shibatani, M. 1999. *The languages of Japan*. Cambridge University Press.
- Sohn, H. 1999. *The Korean language*. Cambridge University Press.
- Tateishi, K. 1990. Phonology of Sino–Japanese morphemes. *University of Massachusetts Occasional Papers in Linguistics* 13, pp. 209–235.
- Tsuchida, A. 1995. English loans in Japanese. *Working Papers of the Cornell Phonetics Laboratory* 10, pp. 145–164.