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): *core–periphery* model of lexicon
- Native vocabulary exists at "core" of lexicon, satisfying all language-specific markedness constraints, while loanwords enter lexicon in outer "periphery" where fewer constraints are obeyed, gradually move inward toward core as they become nativized over time
- Applied to Japanese (Itō and Mester 1995, 1999), Korea (Cho 1999), Québécois French (Paradis and Lebel 1997), American Sign Language (Brentari and Padden 2001)
- Two problems with Japanese data:
  1. No way to consistently rank all markedness constraints governing lexical strata of Japanese
  2. Cannot explain existence of constraints on loanwords not governing other strata

Lexical strata in Japanese

- Four distinct strata
  - Yamato: native vocabulary
  - Sino-Japanese: loanwords from Chinese (7th–13th c.)
  - Mimetic: sound-symbolic vocabulary
  - Foreign: recent loanwords, mainly from English
- Stratal-specific phonotactics (Itō and Mester 1995)
  - *NT: no voiceless stops after nasals (Yamato, Mimetic root-internally)

    - mi**t**a ‘see (past)’; jon**d**a ‘read (past)’
    - Sino-Japanese allows post-nasal voiceless stops (e.g. beN**k**^{j}oː ‘study’)
  - *P: /p/ can only occur geminated or in nasal cluster (Yamato, Sino-Japanese)
    - ni**pp**oN/ni**h**oN ‘Japan’; ya**pp**ari/ya**h**ari ‘after all’
    - i**pp**oN ‘one (bottle, etc.)’; ni**h**oN ‘two (bottles, etc.)’; saN**b**oN ‘three (bottles, etc.)’
    - Mimetic allows word-initial and intervocalic p: **p**ika**p**ika ‘glitter; sparkle’; **p**era**p**era ‘fluent(ly)’
  - Neither *NT nor *P govern Foreign stratum
    - Post-nasal voiceless stops: paŋ**k**u ‘punk’; in**t**aːnetto ‘internet’
    - Word-initial /p/: **p**ɯrosesɯ ‘process’; **p**enparɯ ‘pen pal’
    - Intervocalic /p/: a**p**aːto ‘apartment’; tai**p**ɯ ‘type’

The core–periphery model of the lexicon

- Itō and Mester (1999): w.r.t. *NT and *P, Yamato most constrained, Foreign least constrained, Sino-Japanese in–between
- Lexicon has core–periphery structure in lexical space, with Yamato at core, followed by Sino-Japanese, then Foreign at outermost periphery (mirroring order in which Sino-Japanese and Foreign words were borrowed)
- CODACOND >> FAITH[FOREIGN] >> *P >> FAITH[SINO-JAPANESE] >> *NT >> FAITH[YAMATO]
- Set of possible Yamato words subset of possible Sino-Japanese words; set of possible Sino-Japanese words subset of possible Foreign words
- Foreign not governed by constraints which don’t also govern other strata
- More complex when Mimetic and root–structure constraints taken into account:

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

  - Set of possible Sino-Japanese words cannot be subset of possible Mimetic words, and v.v. (conflicting rankings for *NT and *P)

- ROOT $\leq 2\mu$: root is 1 or 2 morae long, second mora restricted to /i/, /:/, /Q/, /N/, /ki/, /kɯ/, /tʃi/, /tsɯ/ (Tateishi 1990)
- ROOT $= 1\phi$: root is prosodic foot, at least 2 morae (Poser 1990)
- Yamato apparently unrestricted in length (ki 'tree; wood'; me 'eye')

- Core–periphery model not good model of relationship between Yamato, Sino-Japanese, Mimetic
- Core–periphery model predicts that periphery (e.g. Foreign stratum) cannot have unique constraints which don't also govern other strata

Methods

- `JMDICT` (EDRDG 2003); each entry classified according to stratum using computer program
  - Foreign: katakana
  - Sino-Japanese: kanji (on–yomi)
  - Yamato: hiragana, kanji (kun–yomi)
  - Mimetic: bimoraic roots, either reduplicated or with mimetic suffix (/--ri/, /--tosɯrɯ/, etc.); not many listed
  - Hybrid: compounds formed from roots in different strata
  - Hybrid and Mimetic not used in analysis

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	

Comparable to statistics in Shibatani (1999, pp. 142–3)

- Looked for phonotactic patterns rare in Foreign stratum but common in Yamato or Sino-Japanese
- For each candidate pattern, if avoided in loanword adaptation, probably constraint against it (rules out purely static regularities)

Minimal length constraints on loanwords

- English tense vowels borrowed as long vowels ($= 2\mu$), lax vowels borrowed as short vowels ($= 1\mu$) (Lovins 1975, Tsuchida 1995)
- Coda consonants generally disallowed, repaired through epenthesis
- Thus, every loanword (from English) will be $\geq 2\mu$

Open σ, tense V	**Closed σ, lax V**	**Closed σ, tense V**
'key' > kiː	'lip' > rippɯ	'cheap' > tʃiːpɯ
'pay' > peː	'pet' > petto	'cape' > keːpɯ
'show' > ʃoː	'loss' > rosɯ	'rope' > roːpɯ
'you' > yɯː	'put' > pɯtto	'hoop' > ɸɯːpɯ

- Abbreviated loanwords always at least 2 morae long (Itō 1990, Labrune 2002)

2 morae

kʲara	< kʲarakɯtaː	'character'
tʃoko	< tʃokoreːto	'chocolate'

3 morae

arɯmi	< arɯmiɴʲɯːmɯ	'aluminum'
terebi	< terebidʒoɴ	'television'

4 morae

pasokoɴ	< paːsonarɯ+kompʲɯːtaː	'personal computer'
apaːto	< apaːtomento	'apartment'

- Constraint on Foreign stratum:

MINWORD: A word must have at least two morae.

- Three loans found in `JMDICT` which violate MINWORD
  - za 'the' – found only in titles, acts like a bound morpheme
  - ti 'tea', de 'day' – have two–mora variants tiː and deː which are more common (John Whitman, p.c.)
  - Lovins (1975) – one–mora loans short–lived, replaced by two–mora variants created by lengthening vowel
- MINWORD does not govern Sino-Japanese or Yamato (ki 'tree; wood', te 'hand', çi 'sun; sunshine; day', me 'eye')

Palatalized consonants in the Foreign stratum

- Three sources for palatalized consonants:
  1. English /Cju/ > /C^{j}ɯ/ (**d^{j}**ɯːti ‘duty’; bo**r^{j}**ɯːmɯ ‘volume’; koɴ**p^{j}**ɯːtaː ‘computer’)
  2. English /Cæ/ (C a velar stop) > /C^{j}a/ (**k^{j}**atto ‘cat’; sɯ**k^{j}**ana ‘scanner’)
  3. French/Spanish/Italian /ɲV/ > /n^{j}V/ (/ko**n^{j}**akkɯ/ ‘cognac’; /razaː**n^{j}**a/ ‘lasagna’; only about 10 examples in `JMDICT`)
- English /Cjə/ > /CiV/ (V determined by orthography), not /C^{j}V/ (mi**r**ioɴ ‘million’; ita**r**iaɴ ‘Italian’; k^{j}a**n**ioɴ ‘canyon’; ke**n**ia ‘Kenya’)
- Two constraints operating in Foreign stratum:

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

  ***PALLOW:** Palatalized consonants cannot occur before low vowels. (split into two constraints, since velar stops not subject to it)

  (resembles *PALFRONT which holds over entire lexicon)
- MAX–IO[FRONT], DEP–IO, *PALFRONT, *PALMID, *PAL[-VEL]LOW >> CONTIG–IO >> *PAL[+VEL]LOW

  MAX–IO[FRONT] = preserve palatalization on segments (assuming palatalization represented by a feature [+front])

  CONTIG–IO = contiguous segments in output correspond to contiguous segments in input (Kager 1999, p. 250)
- With this ranking, underlying palatalized consonants can only surface before /ɯ/, and violations of *PAL[MID,LOW] will be repaired not through depalatalization or epenthesis, but by creating two segments in the output corresponding to and sharing features of palatalized consonant in input (e.g. /r^{j}/ in input corresponds to /ri/ in output)
  1. ‘value’ > bar^{j}ɯː

/bar^{j}ɯː/	MAX–IO[FRONT]	DEP–IO	*PALMID	CONTIG–IO
☞ [bar^{j}ɯː]				
[bariɯː]				*!
[barɯː]	*!			

2. ‘million’ > mirioɴ, *mirʲoɴ

/mirʲoɴ/	MAX–IO[FRONT]	DEP–IO	*PALMID	CONTIG–IO
[mirʲoɴ]			*!	
☞ [mirioɴ]				*
[miroɴ]	*!			
[mirʲɯoɴ]		*!		*

3. ‘cat’ > kʲatto

/kʲatto/	*PAL[-VEL]LOW	MAX–IO[FRONT]	CONTIG–IO	*PAL[+VEL]LOW
☞ [kʲatto]				*
[kiatto]			*!	
[katto]		*!		

4. ‘Kenya’ > kenia, *kenʲa

/kenʲa/	*PAL[-VEL]LOW	MAX–IO[FRONT]	CONTIG–IO	*PAL[+VEL]LOW
[kenʲa]	*!			
☞ [kenia]			*	
[kena]		*!		

- Sino-Japanese not governed by *PAL[MID,LOW]; palatalized consonants can occur before any non-front vowel (e.g. beɴkʲoː ‘study’)

Conclusion

- Japanese lexicon does not have overall core–periphery structure
  - Yamato, Sino-Japanese, and Mimetic cannot be organized in a subset relationshsip with each other (*NT vs. *P, root structure constraints on Sino-Japanese and Mimetic)
  - Foreign governed by constraints not holding over other strata (MINWORD, *PAL[MID,LOW])
- But core–periphery model does seem to work well for loanword data in other languages: Korean (Cho 1999), Québécois French (Paradis and Lebel 1997), American Sign Language (Brentari and Padden 2001), etc. Korean especially interesting case because similar borrowing history and lexicon structure to Japanese
- In Korean and Canadian dialects of French, loans only make up 1–5% of lexicon (Sohn 1999, Brown 2003), while in Japanese loans make up over 10%

- Core–periphery model seems to be correct for languages in which loans constitute small fraction of total lexicon; but when loans make up 10% or more of lexicon, then perhaps effects of phonotactic constraints on source language become salient in set of loans, loanword constraints derived from them and applied to new loans
- Thus Foreign stratum is not merely a collection of heterogenous loanwords, but seems to be forming a new stratum in Japanese lexicon

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 Conputer 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.