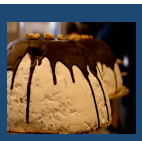
Degemination in Japanese Loanwords from Italian

Maho Morimoto
University of California, Santa Cruz
mamorimo@ucsc.edu



0. Preliminaries

- Italian loans in Japanese: in the musical and culinary domains.
- Host language: Japanese, geminates are distinctive:
  [kata] `shoulder' vs. [katta] `win-PAST'
- Donor language: Italian, geminates are also distinctive:
  [fa:to] `fate' vs. [fatto] `fact'
- Lexical geminates are rare in the native vocabulary, while they are abundant in loanwords.
- I indicate the first part of a geminate with a capital letter.

1. The Puzzle

- When Japanese borrows lexical items from Italian, geminates in Italian can either be preserved…
  espresso /es.préS.so/ → [e.su.pu.réS.so]
  …or be *degeminated*.
  macchiato /maK.kjá:.to/ → [ma.ki.á:.to]
- **What are the patterns? What are the motivations?**
- **How can they be formalized within the framework of Optimality Theory (Prince & Smolensky 1993)?**
- **Are the predictions of the analysis real?**

2. Basic Data & Proposal

Geminate Preservation

- Geminates in Italian borrowings arise as realization of geminates in the source forms (Tanaka 2007):
- Relaxed segmental condition
  farfalle /far.fáL.le/ → [fa.ru.fáL.le]
  glissando /gliS.sán.do/ → [gu.riS.sán.do]
  espresso /es.préS.so/ → [e.su.pu.réS.so]
- These geminates (liquids, [s]) are usually not allowed in other lexical strata.
- Preservation rate (Tanaka's style, my data):

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	-		ɾɾ	17%	(4/23)
tts	53%	(20/38)	ddz	50%	(1/2)			
tʃ	77%	(27/35)	dʒ	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)

Positional Effect

- Geminates in the source forms tend to be preserved in the penultimate syllables in the nativized forms, and they tend to be degeminated outside the last three-syllable window in the nativized forms (Tanaka 2007).
  orecchiette /o.reK.kjéT.te/ → [o.re.ki.éT.te]
  macchiato /maK.kjá:.to/ → [ma.ki.á:.to]
- Preservation rate per position (Tanaka 2007):

further left	4th	antepenultima	penultima	total
29%	38%	60%	73%	60%

- **The last three-syllable window = Italian stress & Japanese pitch accent**

Proposal

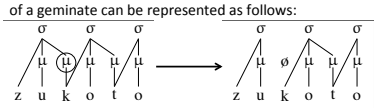
- Relevance of Italian stress:
1. A geminate in a stressed syllable can be preserved when it is not in an accented syllable.
   falsetto /fal.séT.to/ → [fá.ru.seT.to]
   piccolo /píK.ko.lo/ → [piK.ko.ro-]
2. A geminate in a non-stressed syllable can be degeminated even if the mora is in an accented syllable.
   suppli /suP.plí/ → [sú.pu.ri]
   Matteo /maT.téo/ → [má.te.o]

- **Proposal: the preservation of a geminate depends on the head status of the syllable it belongs to in Italian.**

3. A Positional Faithfulness Account

Phonological Representation of Geminates

- Assuming a moraic theory of weight (Hayes 1989), the loss of a geminate can be represented as follows:



O-O Faithfulness in Loanwords

- Assumption: Italian output = Japanese input
- Fully prosodically specified input to Japanese loan phonology:
  zuccotto /zuK.kóT.to/ → [zu.kóT.to]
  orecchiette /o.reK.kjéT.te/ → [o.re.ki.éT.te]

Positional Faithfulness

- The positional effect on degemination can be captured as *stress-based neutralization of consonant length*.
- Positional neutralization ranking schema (Beckman, 1998):
  IDENT-Position[F] » M » IDENT[F]
- A twist: the prominence to which the positional faithfulness depends on can be overwritten by Japanese loanword accent.

Basic OT Analysis

- Constraints:

IDENT-ó[μ]: let β be an input segment in a stressed-syllable, and α its output correspondent. If and only if α is moraic, then β must be moraic.
"An input segment in a stressed syllable and its output correspondent of that segment must have identical moraic specifications."

IDENT[μ]: let β be an input segment and α its output correspondent. If α is moraic, then β must be moraic.
"An input segment and its output correspondent of that segment must have identical moraic specifications."

NoGem: assign a violation for each consonant that is a geminate.

- Relative ranking:
  IDENT-ó[μ] >> NoGem >> IDENT[μ]
- *zuccotto* /zúK.koT.to/ → [zú.koT.to]

/zuK.kóT.to/		IDENT-ˈσ[μ]	NOGEM	IDENT[μ]
☞a.	zu.koT.to		*	*
b.	zuK.koT.to		**!	
c.	zu.ko.to	*!		**
d.	zuK.ko.to	*!	*	*

4. Further Complication

Variability

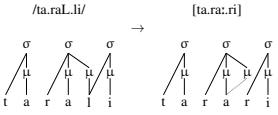
- Losing candidates are attested in free variation:

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

- Account: different adaptation forms belong to different lexical strata (Itô & Mester 1995) with varying rankings of native (NoGem) and loanword-specific constraints (IDENT-ó[μ] >> IDENT[μ]).

Compensatory Lengthening

- Degemination of liquids is sometimes accompanied with a lengthening of the preceding vowel: *taralli*



- This can be accounted for by:

MAX-ó[μ]: assign a violation for each mora in a stressed syllable in the input that is not present in the output.

NoGem[R]: assign a violation for each liquid consonant that is a geminate (after Morén 2001).

MAX[μ]: assign a violation for each mora in the input that is not present in the output.

5. Implicational Hierarchy and A Nonce-Adaptation Survey

Implicational Hierarchy

- Prediction of the Analysis:
1. Preferences among candidates
2. Preferences between *strong* and *weak* geminates
- Degemination patterns and frequency in *tagliatelle:*

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

- What are the actual preferences of Japanese speakers?

Online Survey: Methods

- Online loan adaptation survey using nonce-Italian words.
- Acceptability judgments from 27 native speakers of Japanese, using input & output pairs to rate from 1 to 10.
- Input: 60 three-syllable words containing 2 geminates, varied in types of geminates (liquid vs. voiceless stops).
- Output: 5 possible adaptation patterns in Japanese orthography, varied in operation of geminates (preservation, degemination, compensatory lengthening).

Online Survey: Predictions

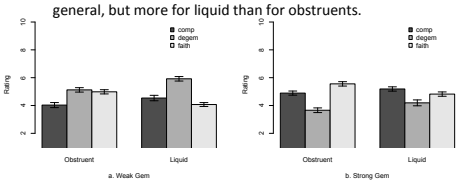
- For an input *bottossa*:
  bo.toS.sa > bo:.toS.sa > bo.to:.sa > boT.to:.sa > boT.to.sa

Online Survey: Sample

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

Online Survey: Results

- Obtained 1620 responses.
- Average rating 4.88.
- Hierarchy revealed ($X^2(1) = 151.04$, $p < 0.001$):
  bo.toS.sa > boT.to:.sa > bo.to:.sa > bo:.toS.sa > boT.to.sa
- Trends confirmed:
1. In strong positions, obstruent geminates prefer to be kept, while liquid geminates prefer not to.
2. In weak positions, geminates prefer to be degeminated in general, but more for liquid than for obstruents.



6. Conclusion

Summary

- The positional effect on degemination in Japanese loanwords from Italian can be captured as stress-based positional neutralization, with the support of stratum-specific rankings of constraints.
- The effect can be formalized using the positional faithfulness schema, assuming an output-output correspondence relationship between the source form and its adapted form.
- Survey results conformed to the predictions except for the implicational relation between faithfulness in strong and weak positions.

Future Work

- Perceptual experiment to test my initial proposal
- Exploration of output-oriented account

References: Beckman, Jill. 1998. *Positional Faithfulness*. PhD dissertation. Hayes, Bruce. 1989. *Compensatory lengthening in moraic phonology*. Linguistic inquiry 253–306. Itô, Junko, and Armin Mester. 1995. The Handbook of Phonological Theory, chapter Japanese Phonology, 817–838. Blackwell. Krämer, Martin. 2009. *The Phonology of Italian*. OUP Oxford. Morén, Bruce. 2001. *Distinctiveness, coercion and sonority: A unified theory of weight*. Psychology Press. Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar*. Tanaka, Shin'ichi. 2007. Itariago no juushiin to sokuon keisei [Geminate consonants in Italian and sokuon in japansese]. *Proceedings of the 134th Meeting of the Linguistic Society of Japan* 252–257.

Acknowledgments: Many thanks to Armin Mester, Junko Itô, Grant McGuire, Pranav Anand, Martin Krämer, members of Winter 2015 Research Seminar, audiences at the Tokyo Circle of Phonologists,, Clara Sherley-Appel, Jeff Adler, Naoya Watabe, Jed Sam Pizarro-Guevara, Allan Schwade, and Mattia Damaggio.
GEMCON 2015 – GLASGOW

Appendices: The Database & Survey Details

Appendix A: The Database

1. Size

- I built a database of Japanese loanwords from Italian, following Tanaka (2007).
- The database contains 1209 Japanized forms total.
- Two different adaptations for a single Italian form are separately counted.
- Entries are concatenated in an Excel spreadsheet with additional information.

2. Sources

- Tokens were hand-picked from 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
- I referred to the etymological information of the dictionaries to
- In order to look up the source word, I used Italian-Italian dictionary, Zingarelli (Zanichelli Editore Spa, 2013).

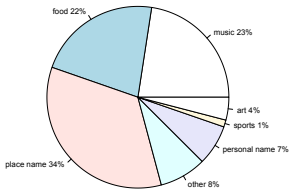
3. Consonants

- Within the 1209 entries,
  - Italian consonants: 5059 occurrences
  - Italian obstruent geminates: 526 occurrences
  - Japanese obstruent geminates: 305 occurrences
- Instances of gemination are quite rare:

rucola	/rú.ko.la/	→	[ruK.ko.ra-]
bufala	/bú.fa.la/	→	[búF.fa.ra]/[buF.fa.ra-]
amatriciana	/a.ma.tri.tʃá:.na/	→	[a.ma.to.riT.tʃá:.na]

4. Domains

- Besides personal names and place names, food and music-related words are prevalent.



Appendix B: Survey Details

1. Survey Materials: the Input

- The input: Italian nonce-words

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

- There were three forms for each type varying in the quality of geminates (liquids vs. stops).

2. Survey Materials: the Output

- 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

- Forms that were rated
  - G, g: voiceless liquid geminates
  - R, r: liquid geminates
  - V, v: compensatory lengthening
  - L, l: light syllable
  - (lower case for weak positions; upper case for strong positions)

	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

3. Survey Materials: the Filler

- There were 60 fillers, taken from Colombo (1992) and Zoccolotti et al. (2005).
- Also trisyllabic Italian nonce-words
- Did not contain any geminates.
- List of 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, onfili, ostura, panchefa, pifato, pirtoci, polaso, poracca, potide, prigiosa, pri- mosta, ravele, rebolo, rudomi, rulate, sintuce, stebore, stevono, stilega, storubo, strebafe, strotula, svepano, tegresto, tirloni, trofulo, trolica, tuposo, vielota, virpico, zerlido

4. Procedure

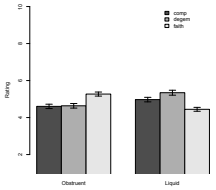
- Participants: 27 native Japanese speakers
- Task: rating the acceptability of 60 critical pairs of input and output in a scale of 1 to 10
- Blocks: 5 blocks with breaks
- Media: Google Form

5. Analysis

- Average rating was 4.88, the most popular responses being 3 and 4 out of 10.
- Responses were analyzed using R (R Core Team, 2013) and *lme4* (Bates et al., 2012).

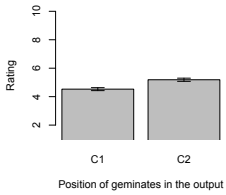
5.1. Effect of Geminate Type

- Different types of geminates prefer different operations to undergo ($X^2(1) = 63.77$, $p < 0.001$).



5. 2. Effect of Geminate Position

- Outputs with geminates in strong positions are more popular ($X^2(1) = 22.26$, $p < 0.001$) .

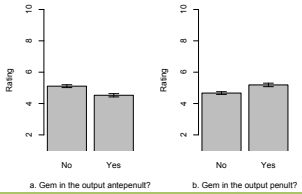


- Trend:

Weak geminates prefer degemination or compensatory lengthening ($X^2(1) = 20.98$, $p < 0.001$)

Strong geminates prefer to be preserved

- Only the former is significant ($X^2(1) = 0.00$, $p < 1$)



Additional references: Bates, Douglas, Martin Maechler, and Ben Bolker. 2012. lme4: Linear mixed-effects models using S4 classes. Colombo, Lucia. 1992. Lexical stress effect and its interaction with frequency in word pronunciation. Journal of Experimental Psychology: Human Perception and Performance 18:987. R Core Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 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.

My paper and this handout are uploaded on my website: http://people.ucsc.edu/~mamorimo