Degemination in Japanese Loanwords from Italian
Maho Morimoto
University of California, Santa Cruz
mmorimoto@ucsc.edu

0. Preliminaries
• Italian loans in Japanese: in the musical and culinary domains.
• Host language: Japanese, geminates are distinctive: [kata] ‘shoulder’ vs. [kata] ‘sheet’.
• Donor language: Italian, geminates are also distinctive: [fa:to] ‘fat’ 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...
express [es.pre:s.so] → [e.su.pu.re.s.so]
or be degeminated.
machìotto [ma:ki.ji.to] → [ma:ki.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 conditional fof forfólate [far.fá.lé] → [fu.ru.fá.lé]
glossando [il.gís.sán.do] → [fu.ru.fís.sán.do]
express [es.pre:s.so] → [e.su.pu.re.s.so]
• These geminates [l]s, [r] are usually not allowed in other lexical strata.

Preservation rate (Tanaka’s style, my data):

<table>
<thead>
<tr>
<th>Type</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>100%</td>
<td>100%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
</tr>
<tr>
<td>II</td>
<td>100%</td>
<td>100%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
</tr>
<tr>
<td>III</td>
<td>100%</td>
<td>100%</td>
<td>97%</td>
<td>97%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
</tr>
<tr>
<td>IV</td>
<td>100%</td>
<td>100%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
</tr>
<tr>
<td>V</td>
<td>100%</td>
<td>100%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
<td>97%</td>
</tr>
</tbody>
</table>

• Preservation rate per position (Tanaka 2007):

<table>
<thead>
<tr>
<th>Position</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>2nd</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>3rd</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>4th</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>5th</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>6th</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>7th</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>7th</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

...is known as phonological adaptation.

Proposal
• Relevance of Italian stress:
  1. A geminate in a stressed syllable can be preserved when it is not in an accented syllable.
  2. A geminate in a non-stressed syllable can be degeminated even if the mora is in an accented syllable.

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:

Basic OT Analysis
Constraints:
IDENT-d{l}: 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[β]: 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-d{l} >> NoGem >> IDENT[β]

<table>
<thead>
<tr>
<th>NoGem</th>
<th>IDENT[β]</th>
<th>IDENT-d{l}</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>70%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Variability
• Losing candidates are attested in free variation:

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>[takii]</td>
<td>10%</td>
</tr>
<tr>
<td>[takō]</td>
<td>90%</td>
</tr>
</tbody>
</table>

• Account: different adaptation forms belong to different lexic truncation (Itō & Mester 1995) with varying rankings of native (NoGem) and loanword-specific constraints (IDENT-d{l} >> IDENT[β]).

Compensatory Lengthening
• Degemination of liquids is sometimes accompanied with a lengthening of the preceding vowel: tanoll [ta:noll] [ta.noll]

This can be accounted for by:
MAX-d{l}: 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 Morin 2001).
MAX-A: assign a violation for each mora in the input that is not present in the output.

4. Further Complication

Summary
• The positional effect on degemination in Japanese loanwords from Italian can be formalized 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


Acknowledgments: Many thanks to Armin Mester, Junji Kii, Grant McGauley, Pasqua Andron, Martin Krämer, members of Winter 2015 Research Seminar, audiences at the Tokyo Circle of Phonologists, Chris Shin, Akihito Abe, Jeff Keller, Yuji Yotaka, Jiri Pluta, Vincenzo Ascoli, Allen Schade, and Mattia Dorigatti.

0EMCON 2015 – GACGIRB

Heroes: Hana Morimoto for helping me to write this. I am in desperate need of her help. This is, however, the easiest piece I have written in recent times.
Appendix: 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:
  Kūjen, Shimura 1998
  Japanese pronunciation accent dictionary, NHK Höō 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 obstructive geminates: 526 occurrences
  Japanese obstructive geminates: 305 occurrences
- Instances of gemination are quite rare:
  rucola /ru.ka.ro/ → [ruK.ta.ro-]
  bufala /bu.ca.ro/ → [buF.ta.ro]/[buF.ta.ro-]
  amariciana /a.ma.tri.tʃ:na/ → [a.ma.to.rI.tʃ: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

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>gGI</td>
<td>ciuffocco</td>
</tr>
<tr>
<td>gRI</td>
<td>gorruppa</td>
</tr>
<tr>
<td>rGI</td>
<td>collerre</td>
</tr>
<tr>
<td>rRI</td>
<td>ciollerre</td>
</tr>
</tbody>
</table>
- 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

<table>
<thead>
<tr>
<th>Output example</th>
<th>weak gem</th>
<th>strong gem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. gGl bo.to.so.sa</td>
<td>faith</td>
<td>faith</td>
</tr>
<tr>
<td>2. gV1 bo.to.so.sa</td>
<td>faith</td>
<td>comp</td>
</tr>
<tr>
<td>3. gL1 bo.to.so.sa</td>
<td>faith</td>
<td>degem</td>
</tr>
<tr>
<td>4. vGl bo.to.so.sa</td>
<td>comp</td>
<td>faith</td>
</tr>
<tr>
<td>5. vV1 bo.to.so.sa</td>
<td>comp</td>
<td>comp</td>
</tr>
<tr>
<td>6. vL1 bo.to.so.sa</td>
<td>comp</td>
<td>degem</td>
</tr>
<tr>
<td>7. lGl bo.to.so.sa</td>
<td>degem</td>
<td>comp</td>
</tr>
<tr>
<td>8. lV1 bo.to.so.sa</td>
<td>degem</td>
<td>degem</td>
</tr>
<tr>
<td>9. lL1 bo.to.so.sa</td>
<td>degem</td>
<td>degem</td>
</tr>
</tbody>
</table>

- 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)

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: batto, bildeze, birfola, bitona, blosidi, bolitic, bortaca, borteso, botume, canfrasto, cegape, celimo, cirtora, dilone, dinuro, drivule, fanziane, fastanda, flenesta, fromile, grocelo, iselo, laromo, linebre, lintere, livero, loriaia, marlipo, meribe, mevino, olina, onfill, ostura, panchefa, pifato, pictori, polaso, poracca, potide, prigiosa, pri-mosta, ravele, rebolo, rudomi, rulate, sintuce, stebore, stevono, stilega, storugo, strebafe, strotula, svepano, tegresto, tiri Toni, trofulo, trolica, tuposo, violota, 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 Ime4 (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 ($X^{2}(1) = 0.00, p < 1$).


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