Loanword adaptation is not just perception:
Evidence from Japanese loan doublets

Jennifer L. Smith
University of North Carolina, Chapel Hill
jlsmith@unc.edu

I. Background: The significance of loanword adaptation

(1) Loanword
   (a) A word from one language — the source language (L_s)
       ...is used by speakers of another — the borrowing language (L_b)
   (b) Example:  L_s/English    cream    [kju:m]
                 L_b/Japanese  クリーム    [ku:ri:mu:u] 'cream'

(2) Loanword adaptation:  Phonological alterations that make a loanword
                          compatible with the phonology of L_b

(3) Example of loanword adaptation:
    • Eng (L_s) cream [kju:m] —> violates several Jpn (L_b) phonotactic constraints
    (a) Onset clusters not permitted:  *[kr]
    (b) Non-geminate, non-/N/ codas not permitted:  *[m]
    (c) In the adapted L_b form, epenthesized [u]s avoid these problems
        ▶ [ku:ri:mu:u]

(4) Questions to ask about loanword adaptation
   (a) How are loanwords adapted?
       • When more than one repair is possible, which one is chosen?
         L_s [kju:m] —> L_b [ku:ri:mu:u], not *[_ri:_]
   (b) Where (in the language system) are loanwords adapted?
       • Does adaptation bear on speech perception?  L_b phonology?  UG?

(5) Outline of the talk
   § II  Background on traditional L_b-phonology view of loanword adaptation
   § III Alternative:  loanword adaptation as perception (Peperkamp & Dupoux 2003)
   § IV  Critique of perception-only view:  loan doublets in Japanese
   § V  Alternative proposal:  OO-FAITH in loanword phonology (summary today)
II. 1st approach: Loanword adaptation is Lb phonology


(a) The underlying representation (UR) in Lb closely resembles the Ls form (at least for speakers who actually borrow a word, in contact with Ls)

(b) The phonological grammar of Lb then maps that UR to a surface representation (SR) that conforms to Lb phonology

(7) Example to illustrate this view: Ls [kijim] —> Lb [kuiri:mu]

(a) In Japanese, the UR of borrowed cream is (was) /kri:m/

(b) This UR undergoes epenthesis to produce the SR [ku:ri:mu]

<table>
<thead>
<tr>
<th>/kri:m/</th>
<th>NO COMPLEX ONSET</th>
<th>CODA CONDITION</th>
<th>NO DELETION</th>
<th>NO EPENTHESES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. kri:m</td>
<td>*(!)</td>
<td>*(!)</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>
| 2. ku:ri:mu | | | | *!*
| 3. _ri:_ | | | | |

Constraint definitions

- NOCOMPLEXONSET: Onset clusters are prohibited (Prince & Smolensky 1993)
- CODACONDITION: Codas with non-shared coda Place features are prohibited (Itô 1989)
- NOEPENTHESES (DEP): Output segments have input correspondents (McCarthy & Prince 1995)
- NODELETION (MAX): Input segments have output correspondents (McCarthy & Prince 1995)

(c) A crucial aspect of this analysis: NODELETION >> NOEPENTHESES

This ranking is what makes epenthesis the preferred repair

(8) Potential difficulty for the account: In Japanese non-loan phonology, the default repair is arguably deletion (McCawley 1968)

- Vowel-final verbs: Suffixes surface unchanged
- Consonant-final verbs: Suffix consonants delete

<table>
<thead>
<tr>
<th>nonpast /-ru:/</th>
<th>causative /-sase/</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) 'see' /mi-ru/</td>
<td>[mi.ru] /mi-sase/ [mi.sa.se]</td>
</tr>
<tr>
<td>(b) 'eat' /ta-be-ru/</td>
<td>[ta.be.ru] /ta-be-sase/ [ta.be.sa.se]</td>
</tr>
<tr>
<td>(c) 'read' /jom-ru/</td>
<td>[jom.u] /jom-sase/ [jom.a.se]</td>
</tr>
<tr>
<td>(d) 'fly' /to-b-ru/</td>
<td>[to.b.u] /to-b-sase/ [to.b.a.se]</td>
</tr>
</tbody>
</table>

- Opposite ranking is motivated here: NOEPENTHESES >> NODELETION
This is an example of a more general research problem: Why is there an *epenthesis preference* in loanword adaptation?

(a) Beyond Japanese, there is a *cross-linguistic tendency to avoid deletion* when adapting loanwords (Paradis & LaCharité 1997)

(b) Most extreme consequence: the languages that

- use deletion as the default repair strategy in non-loan phonology
  - have the constraint ranking NOEPENTHESIS >> NODELETION

- BUT choose epenthesis repairs in loanword adaptation

(c) Languages with epenthesis only for loanwords show that the non-loan phonology cannot be the only mechanism in loanword adaptation

### III. New approach: Loanword adaptation is perception

(10) Alternative view: Loanword adaptation is entirely a perceptual effect (Peperkamp & Dupoux 2003, Peperkamp to appear)

(a) Psycholinguistic experiments show that speech perception can be distorted by the listener's native phonological system (e.g., Best 1994, Hallé et al. 1998, Dupoux et al. 1999, Moreton & Amano 1999, Mielke 2003)

(b) Peperkamp & Dupoux's proposal

- All adaptation occurs during perception
- Lₕ phonology influences how Lₛ words are (mis)perceived, but there is no UR--SR mapping in loanword adaptation

(11) Example to illustrate this view

(a) Lₛ acoustic signal (English *cream*)

(b) Lₕ speaker's **phonetic decoding module** maps the acoustic signal to a surface phonetic representation

(c) Lₕ speaker's **phonological decoding module** maps the surface phonetic representation onto a corresponding underlying representation
(12) Phonetic decoding in the P&D model (emphasis added)

(a) The phonetic decoding module maps acoustic information onto the closest native phonetic category

(b) "With respect to nonnative sounds, this mapping is of course massively unfaithful, since the phonetic categories to which these sounds are mapped in the foreign language can simply be absent from the native one." (Peperkamp & Dupoux 2003: 368)

(c) "The phonetic decoder takes not only the inventory of segments into account but also those of suprasegments and syllables. Hence, nonnative suprasegments and syllable types are mapped onto the closest native ones." (Peperkamp & Dupoux 2003: 369)

(13) Prediction of the P&D model for Japanese:
Epenthesis in Japanese loanword adaptation is a consequence of perception.

(14) So, is Japanese loanword epenthesis solely a perceptual effect?

(a) Phonological reasoning: If loanword epenthesis is not a product of the UR→SR mapping, there is no contradiction between non-loan phonology (deletion) and loanword adaptation (epenthesis).
• Peperkamp (to appear) makes an argument like this for similar cases.

(b) Psycholinguistic evidence: Japanese listeners have more difficulty than French listeners in distinguishing auditory stimuli with ...CC... and ...CuC... (Dupoux et al. 1999; Dehaene-Lambertz et al. 2000)

• They propose that this is because both CC and CuC forms are mapped onto CuC by the Japanese phonetic decoding module. That is, CC is perceived as CuC: "perceptual epenthesis"

(c) Outstanding questions:
• Is perceptual epenthesis consistent in the medial ..CC.. environment?
• Can perceptual epenthesis be consistently demonstrated for Japanese listeners in other phonological environments?

(15) Counterproposal — Today's focus is on part (a)

(a) Additional perceptual evidence: Adaptation patterns in auditory loanwords suggest that epenthesis is not an automatic consequence of perception by Japanese speakers of Ls forms with "problem" consonants

(b) Phonological reasoning: Loanword adaptation and non-loan phonology can differ because only loanwords involve OO-FAITH to the Ls form
IV. Evidence from Japanese loan doublets

(16) Empirical argument: Peperkamp & Dupoux's (2003) claim is too strong

- Not all perceptual effects lead to epenthesis in Japanese

Consequence: A UR→ SR mapping is involved in loanword adaptation

(17) Overview of data in (18)-(20)

- Words borrowed in ways unlikely to be influenced by L’s orthography sometimes show consonant deletion instead of vowel epenthesis

(18) English phrasebooks for 19th-century merchants (Kamei et al. 1965)

(a) Final coda deletion

- Item: ワリワン [wa.ri.wan_] K:147
- Gloss: nan de gozaru ‘what is it?’
- Probable source: < what [do] you want

- Item: ナイ [nai_] K:148, from Nihon gaikoku shounin dokutsuushi
- Gloss: yoru ‘evening, night’
- Probable source: < night; cf. [nai.to] (Arakawa 1977)

(b) Medial coda (geminate) simplification by deletion

- Gloss: okoru ‘become angry’

(19) Plausibly auditory loanwords, 19th-20th century

- Data from Arakawa 1977, Ichikawa 1929, Miura 1993

Notes:
- Doublet forms given in the rightmost column are either from the same source as the “non-conventional” form, or else from Arakawa (1977)
- The symbol ‘<’ indicates a deletion form that only differs from its epenthesis doublet form in having the shape [(µµ µµ)] or [(µµ µµ)], and moreover matches the epenthesis form at the left edge(s) of the L’s morpheme(s). A case like this may have an alternative analysis as a metrically motivated truncation of the epenthesis form (see Itô 1990 on the phonology of loan truncation).
- Exception: If a deletion form that meets these criteria has different representations of L’s vowels than the epenthesis form (especially L’s reduced vowels), this is taken as evidence of auditory borrowing rather than metrically motivated truncation.

(a) Onset cluster simplification by deletion
doublet form with epenthesis

- [ri.su.rin] glycerine ‘glycerine’ I:25 [gu.ri.se.rin]
- [wai.ja.tu] white shirt ‘white/dress shirt’ I:8 [ho.wai.to]

(b) Final coda deletion
doublet form with epenthesis

- [i.ru.ba] jitterbug ‘jitterbug’ A:577 [dit.ta.bag.gu]
- [pok.ke] pocket ‘pocket’ I:7 [po.ket.to]
- [ra.mu.ne] lemonade ‘lemon-flavored drink’ I:3, M:171 [re.mo.ne.do]
- [han.ket] handkerchief ‘handkerchief’ I:7, M:136 [han.ca.i.:fu]
- [o.rai] all right ‘all right’ I:32 [o.:ru rai.to]
(c) Final coda-cluster simplification by deletion
doublet form with epenthesis

\[\text{crank} \rightarrow \text{crank'} I:26\]
\[\text{cement} \rightarrow \text{cement'} I:26\]
\[\text{tongue} \rightarrow \text{tongue (food')} I:4, M:177\]
\[\text{beefsteak} \rightarrow \text{beefsteak'} I:2\]
\[\text{next year} \rightarrow \text{generation}, \text{next year'} I:32\]
\[\text{all right} \rightarrow \text{all right'} I:32\]

(d) Coda \[\eta\] as \[N\], not \[N\]
doublet form with epenthesis

\[\text{pudding} \rightarrow \text{pudding'} I:3\]
\[\text{tongue} \rightarrow \text{tongue (food')} I:4, M:177\]
\[\text{surfing} \rightarrow \text{surfing'} M:139\]

(e) Medial coda deletion
doublet form with epenthesis

\[\text{Hepburn} \rightarrow \text{Hepburn'} I:58\]
\[\text{white shirt} \rightarrow \text{white/dress shirt'} I:8\]
\[\text{beefsteak} \rightarrow \text{beefsteak'} I:2\]
\[\text{next year} \rightarrow \text{generation}, \text{next year'} I:32\]
\[\text{all right} \rightarrow \text{all right'} I:32\]

(20) \textbf{English loanwords in Hawaiian Japanese} (Higa 1970)

- Open question: To what extent have these loanwords been influenced by Hawaiian Creole English ("Hawaiian Pidgin"), which also has deletion repairs?
- Possible metrically motivated truncations (see (19)) indicated with '•'

(a) Deletion of final voiced stop, \(V\)\# (H:137)

\[\text{inside} \rightarrow \text{inside'} I:8\]
\[\text{outside} \rightarrow \text{outside'} I:2\]

(b) Deletion of final voiceless stop, \(N\)\# (H:131)

\[\text{husband} \rightarrow \text{husband'} I:12\]

(c) Deletion of voiceless stop, \(S\)\# (H:136)

\[\text{next year} \rightarrow \text{next year'} I:32\]
\[\text{last year} \rightarrow \text{last year'} I:32\]

(d) Deletion of medial-coda voiceless stop (H:137)

\[\text{outside} \rightarrow \text{outside'} I:32\]

(21) Conclusion to be drawn from the loan-doublet data:

 Auditory perception of English forms by Japanese speakers does not always lead to epenthesis (although it may in some instances or certain contexts)
**V. A phonological account for Japanese loanword adaptation**

(22) If "perceptual deletion" can occur when Japanese speakers perceive Ls words, why is epenthesis as such a pervasive strategy in loanword adaptation?

(a) Many loanwords have come from written materials (see, e.g., Miura 1993)

(b) Consonants that might not be perceived in auditory input are accessible from the orthographic forms

(23) Deriving the doublet from Ls *jitterbug*

(a) **Auditory borrowing**

(assuming the Peperkamp & Dupoux 2003 model of auditory perception)

i. Ls phonetic form [dʒi."t.b'j"]

ii. Acoustic form

iii. Lb phonetic decoding module [dʒi.ru.b_] → perception may lead to deletion

iv. Lb underlying form /di.ru.ba/

v. Lb surface form [dʒi.ru.ba]

(b) **Orthographic borrowing**

i. Ls spelling <jitterbug>

ii. Assumed target form → UR [dʒit.ta.bag(g)]

(on the basis of orthographic decoding)

iii. Lb surface form [dʒit.ta.bag.ɡɯ] → phonological epenthesis

▷ [ɡ] in Ls form is "perceived" via orthography

▷ However, [u] in adapted form is not provided by orthography

▷ This epenthetic vowel is the result of a UR→SR mapping

(24) Why deletion in non-loan phonology, but epenthesis in loanword adaptation?

(a) Loanword phonology in Lb often includes a drive to match the (perceived) Ls form — but this is a matter of *violable* constraints, rather than being a mandatory outcome of speech perception

▷ For evidence that perceptual similarity in loanword adaptation is violable and interacts with other aspects of the phonological grammar, see (e.g.) Davidson & Noyer 1997, Steriade 2001, Yip 2002, Kang 2003.

(b) This drive to match the Ls form with loanwords in Lb can be instantiated as a set of correspondence (faithfulness) constraints on the Ls-O dimension, a type of output-output faithfulness where the 'base' is the source-language pronunciation (or the orthographic representation thereof) → similar to Yip's (2002) MIMIC constraints
(25) Other problems for a perception-only view of loanword adaptation and the epenthesis preference

(a) **Creoles** use both epenthesis and deletion repairs (Alber & Plag 2001)

(b) Some languages maintain **non-native structures** in loanwords

(c) "Perceptual assimilation" is **not exceptionless** even for non-native structures (see, e.g., Hallé et al. 1998, Mielke 2003)

(26) **Conclusions**

- Loanword perception is certainly influenced by native-language phonology
- But, "perceptual assimilation" is not the only force in loanword adaptation
- A phonological analysis that includes **OO-FATH** between Lb forms (as perceived by Lb speakers) and Lb forms is best able to capture the diverse forces that influence loanword phonology — including not only perceptual effects, but also orthography and interactions between loanword adaptation and other phonological constraints active in Lb

**References**


