24.901 Loanword Phonology
Dec.6, 2010

1. motivations for study of loanword adaptation
   • test productivity of rules and constraints
   • with shift to constraint-based grammar can better model conflict between
     faithfulness to the source language and conformity with the segmental inventory,
     phonotactics, and prosodic structures of the recipient language
   • English flight [flæt] > Korean [pʰʌ.la.tʰi]
   • "static" languages like Thai and Chinese figure prominently
   • learnability puzzles: how are adaptation strategies learned since they may be
     absent from the native grammar?

2. terminology

   donor language  recipient language
   L2               L1
   source word     adapted word

   import: novel sound or structure introduced via loans

3. adaptations
   • typically sharp intuitions
   • robust and productive (hundreds or thousands of loanwords in Japanese, Korean)
   • conservative: if match is available then it is typically made
   • grammatical calculation based on some underlying competence
   • what formal properties?: input, output, internal states

4. major questions in recent literature

What is the input to adaptations process? Some answers offered
   • acoustic signal of L2 output: extragrammatical speech perception module filters
     speech at the outset
   • an accurate IPA like transcription of L2 output (some degree of bilingualism)
     which is then filtered by L1 grammar
   • abstract phonological representation of loan: bilinguals draw on phonological
     hearing in both L1 and L2 and ignore the details of phonetic implementation
   • relative contribution of oral vs. orthographic transmission

Does loanword adaptation implicate a special grammar distinct from L1 native
grammar? Universal principles of sound perception and similarity?
   • if grammar is just rewrite rules then answer must be yes since no evidence for
     epenthesis process in Cantonese native grammar where English bus is adapted
     as [pasi]
12. Input to loanword adaptation process
- phonemic representation that abstracts away from predictable phonetic variants in donor language (Paradis & Lacharité 2005)
- based on phonological hearing by bilinguals
- alternative adaptation strategy based on phonetic salience (Kenstowicz 2003)

13. Mandarin case study (Hsieh, Kenstowicz, & Mou 2006)
- vowel phonemes: i y u
- a
- coda: glides j and w and nasals n and n
- mid and low vowels have various realizations depending on coda (and onset)

\[\overline{o}\] back unround [ɔ] in open syllable: kɔ [kʰɔ] 4 'guest'
back round [o] adjacent to [w]: xɔw [xɔw] 4 'thick'
front [ɛ] adjacent to [j]: kj [kʰɛj] 1 'scold'
central [ə] elsewhere: /ken/ [kʰən] 3 'willing'

\[\overline{a}\] back [a] before n and w: kang [kʰəŋ] 2 'carry on shoulder', kaw [kʰəw] 3 'test'
front [ɛ] when preceded by [j] and followed by [n]: jian [tsjɛn] 4 'build'
elsewhere central-front [a]: da [ta] 4 'big'

14. prediction: if loanword adaptation abstracts away from noncontrastive details then we expect the following mappings based on phonological vs. phonetic input models:

**phonological mapping**

<table>
<thead>
<tr>
<th>English</th>
<th>æn</th>
<th>æŋ</th>
<th>an</th>
<th>æŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>an</td>
<td>Mandarin</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>æŋ</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>an</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>æŋ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

in phonetic model either then vowel or the nasal could determine the adaptation; perhaps preference to the vowel as the more dominant perceptually

**phonetic mapping**

<table>
<thead>
<tr>
<th>English</th>
<th>æn</th>
<th>æŋ</th>
<th>an</th>
<th>æŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>an</td>
<td>Mandarin</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>æŋ</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>an</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>æŋ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. loans
  • c. 600 VN loans from *Dictionary of Loanwords and Hybrid Words in Chinese*.
  • place names from Chinese Ministry of Foreign Affairs website

5. data

i. dictionary
  • [æn] -> [an] (26): an.chou 'anchovy'; ke.lan 'clan'
  • [æn] -> [an] (5): wang.da.'er 'vandel'
  • [æŋ] -> [əŋ] (5): gang.guo 'Congo'
  • [æŋ] -> [əŋ] (2): an.ge 'encore'
  • [əŋ] -> [əŋ] (24): a.nang 'anonymous'; kang.tuo 'canto'
  • [əŋ] -> [əŋ] (0)
  • [æŋ] -> [əŋ] (9): fulan 'Franklin'
  • [æŋ] -> [əŋ] (4): jang.ji 'Yankee'

ii. place names
  • [æn] -> [an] (21): Manhattan man.ha.dun
  • [æn] -> [ən] (0)
  • [ən] -> [æŋ] (15): Oregon e.ιe.ɡang
  • [ən] -> [ən] (3): Tucson tu.sen
  • [æŋ] -> [əŋ] (1): Doncaster tang.ke.si.te

iii. personal names: Lin (2006)

Johnson qiāng.shēng [tʃhjoŋ.ʂəŋ] 2.1

6. conclusion
  • adaptation is based on phonologically redundant but phonetically salient [±back]
  vowel difference that accompanies and enhances the coronal vs. velar coda
  contrast
  • place contrasts among coda nasals are often difficult to hear
5. case study of loans from English to Cantonese and Mandarin

- not as many as Japanese or Korean
- written vs. oral transmission


\[
\begin{array}{cccccc}
p & t & t^h & k & k^w \\
p^h & t^h & t^{h} & k^h & k^{wh} \\
f & s & h \\
m & n & \eta \\
y & w \\
\end{array}
\]

syllable: CVC

tones: 55 35  
33 24  
22 21

7. Matching segments correspond

fun fAn 55  
cut kAt55  
tie tha:i 55

8. closest matching segment

- \( r > 1 \): rum lAm 55
• $f > s$: show sou55
• $tf > ts$: check t'ek 55

8. novel syllables created (cf. Mandarin)
• accidental gaps in onset+rhyme: fit $>$ fit55, win $>$ win
• new rhymes: band $>$ pen, game $>$ kem
• violate ban on labials: pump $>$ pAm

9. final consonants

Stops: sharp [sap]
cut [kʰAt]
Jack [tsik]

Nasals: jam [tsem]
gin [tstin]
bowling [pouliŋ]

s:
lace [leisi]
office [ofisi]

custers:
band [pen] shaft [sep] notes [nuksi]
friend [fen] post [p'ousi] tips [tipsi]
pump [pem] cast [k'asi] licence [laismen]
stamp [sitam] toast [tɔsi] inch [intsí]
sink [siŋ] waist [wesi] yeast [isi]
foul [fAu] fight [fAi]

10. initial cluster reduction
• disyllabic minimal word preference
• no complex onsets or codas
• sC clusters repaired by epenthesis while CR repaired by deletion of liquid--a contrast in contextual saliency
• but, liquid is retained if output would fall below two syllables (Silverman)
• data from Chan & Kwok '82 and other sources

<table>
<thead>
<tr>
<th>CR deleted</th>
<th>CR preserved</th>
</tr>
</thead>
<tbody>
<tr>
<td>printer</td>
<td>brake</td>
</tr>
<tr>
<td>p'en.t'a S 92</td>
<td>pik.lik</td>
</tr>
<tr>
<td>broker</td>
<td>cream</td>
</tr>
<tr>
<td>puk.k'a</td>
<td>kei.lim</td>
</tr>
<tr>
<td>floorshow</td>
<td>fluke</td>
</tr>
<tr>
<td>fosow</td>
<td>fu.luk</td>
</tr>
<tr>
<td>freezer</td>
<td>clean</td>
</tr>
<tr>
<td>fi.sa</td>
<td>ki.lin S 92</td>
</tr>
<tr>
<td>place</td>
<td>flea</td>
</tr>
<tr>
<td>p'ei.si</td>
<td>fu.li S 92</td>
</tr>
<tr>
<td>professor</td>
<td>blonde</td>
</tr>
<tr>
<td>pou.fa.sa</td>
<td>pi.lan S 92</td>
</tr>
<tr>
<td>high-class</td>
<td>pleat</td>
</tr>
<tr>
<td>hai.k'a.si</td>
<td>p'i.lit S 92</td>
</tr>
<tr>
<td>blender</td>
<td>plum</td>
</tr>
<tr>
<td>p'en.ta S '92</td>
<td>pow.lamY '93</td>
</tr>
<tr>
<td>strawberry</td>
<td>print</td>
</tr>
<tr>
<td>sitawpeley Y '93</td>
<td>p'i.lin S '92</td>
</tr>
</tbody>
</table>
proton pow-ton S 92
price p'ay-si S 92

text in red:

exceptions

friend fæn
dacron
gross lo
brandy
clutch
petlan.tei
kiklik.tsı

- contrast between fluke -> fuluk vs. place -> p'ei.sı or blonde -> pi.lan vs. blender
  -> p'en.ta is striking. Why should deletion of [l] depend on whether or not
  the word ends in [s]? (cf. OT's "global comparison")
- assuming that simplification of CR clusters eliminates the minimally salient R, this
  process must be blocked until the disyllabic phonotactic from the Operative stage
  is imposed--inconsistent with Silverman's model.

- syllable with major stress in English assigned 55
- final unstressed varies between 21 and 35 (latter reflects a floating H suffix)
- pretonic has 22

examples

- monosyllables: 63 of 70 monosyllables take 55
  
card khát 55
  lift lip 55
  tie tha:i 55

- SW disyllables
  
boxing pok 55 siŋ 35
party phat 55 ti 21
partner pa:t 55 la: 35/21

- WS
cologne kɔ 22 luŋ 55

- epenthetic vowels

  bus pa: 55 si 35
  cash khe: 55 sy 21
  cream kei 22 lim 55

  stamp sì22 ta:m 55
  store sì22 tɔ 55
  custard kAt55sì22ta35
Cantonese vowel system (Bauer 1997)

<table>
<thead>
<tr>
<th>high</th>
<th>i:</th>
<th>y:</th>
<th>u:</th>
</tr>
</thead>
<tbody>
<tr>
<td>closed</td>
<td>e</td>
<td>ø</td>
<td>o</td>
</tr>
<tr>
<td>open</td>
<td>e:</td>
<td>æ:</td>
<td>ə:</td>
</tr>
<tr>
<td>central</td>
<td>ə</td>
<td></td>
<td></td>
</tr>
<tr>
<td>glide</td>
<td>j</td>
<td>a:</td>
<td>w</td>
</tr>
</tbody>
</table>

- no long/short contrast independent of timbre; is basic contrast one of quality or quantity?
- long vowels outnumber short
- short vowels barred from open syllable: *CV (bimoraic)

Rhyme table (Bauer 1997)

```
<table>
<thead>
<tr>
<th>29</th>
<th>56 Cantonese rimes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>j (iyorw)</td>
<td>m</td>
</tr>
<tr>
<td>[i]</td>
<td>iorw</td>
</tr>
<tr>
<td>[o]</td>
<td>—</td>
</tr>
<tr>
<td>[u]</td>
<td>—</td>
</tr>
<tr>
<td>[ə]</td>
<td>—</td>
</tr>
<tr>
<td>[e]</td>
<td>eə</td>
</tr>
<tr>
<td>[e]</td>
<td>—</td>
</tr>
<tr>
<td>[i]</td>
<td>—</td>
</tr>
<tr>
<td>[o]</td>
<td>—</td>
</tr>
<tr>
<td>[u]</td>
<td>—</td>
</tr>
<tr>
<td>[ə]</td>
<td>—</td>
</tr>
<tr>
<td>[e]</td>
<td>eə</td>
</tr>
<tr>
<td>[e]</td>
<td>—</td>
</tr>
<tr>
<td>[i]</td>
<td>—</td>
</tr>
<tr>
<td>[o]</td>
<td>—</td>
</tr>
<tr>
<td>[u]</td>
<td>—</td>
</tr>
<tr>
<td>s</td>
<td>—</td>
</tr>
</tbody>
</table>

* Asterisked vowels in the above table do not occur as independent rimes in open syllables.
```
Gaps in core vocabulary

- round vowels do not combine with labial consonants: *[-cons, +labial] [+cons, +labial]
- front mid vowels do not combine with labials or dentals: *[-high, -low, -back] [+cons, +anterior]
- high vowels do not combine with velars while closed mid vowels are only found before velars and so closed mid vowels are usually treated as allophones of high vowels
- constraints:
  a). OK: * [-cons, +labial] [+cons, +labial]
  (b). *ET: *[-high, -low, -back] [+cons, +anterior]
  (c). *[ + high] [+cons, + high] >> *[-high, -low, -back, +ATR]

Base-line vowel correspondences-1 Open syllable

<table>
<thead>
<tr>
<th>English</th>
<th>Cantonese</th>
</tr>
</thead>
<tbody>
<tr>
<td>[iː]</td>
<td>CD</td>
</tr>
<tr>
<td>[uː]</td>
<td>boot</td>
</tr>
<tr>
<td>[eː]</td>
<td>ATM</td>
</tr>
<tr>
<td></td>
<td>gay</td>
</tr>
<tr>
<td>[oː]</td>
<td>OK</td>
</tr>
<tr>
<td>[aː]</td>
<td>bar</td>
</tr>
<tr>
<td>[æ]</td>
<td>IQ</td>
</tr>
<tr>
<td></td>
<td>high</td>
</tr>
<tr>
<td></td>
<td>pie</td>
</tr>
<tr>
<td>[æw]</td>
<td>foul</td>
</tr>
</tbody>
</table>

Base-line vowel correspondences-2 Closed syllable

<table>
<thead>
<tr>
<th>English</th>
<th>Cantonese</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i]</td>
<td>kid</td>
</tr>
<tr>
<td></td>
<td>pin</td>
</tr>
<tr>
<td></td>
<td>lift</td>
</tr>
<tr>
<td>[ɛ]</td>
<td>chemistry</td>
</tr>
<tr>
<td></td>
<td>sex</td>
</tr>
<tr>
<td></td>
<td>to check</td>
</tr>
<tr>
<td></td>
<td>set</td>
</tr>
<tr>
<td>[ʌ]</td>
<td>fun</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>Cantonese</th>
</tr>
</thead>
<tbody>
<tr>
<td>[iː]</td>
<td>kid</td>
</tr>
<tr>
<td></td>
<td>pin</td>
</tr>
<tr>
<td></td>
<td>lift</td>
</tr>
<tr>
<td>[ɛː]</td>
<td>chemistry</td>
</tr>
<tr>
<td></td>
<td>sex</td>
</tr>
<tr>
<td></td>
<td>to check</td>
</tr>
<tr>
<td></td>
<td>set</td>
</tr>
<tr>
<td>[ʊ]</td>
<td>fun</td>
</tr>
</tbody>
</table>
Matching vowel quality over length

/\c/ck/  
\>tsh:ck  *
\tsh:ck  *
\tsek  *

Phonotactic Constraint-1: *ET

a.  [e]  cents  [i:]  si:nsi:
pence  pi:nsi:  22-22
cassette  [e]  ka:sek
cheque  tsek  C&K
offset  o:sek
b.  [ae]  jam  tsi:m  C&K
Benz  [e:]  pe:nsi:
to check  tsek  C&K
cancer  [e:]  khë:nsá:
jam  tsë:m
band  pë:n

/p[e]nce/  *ET  OK  Id-[high]  Id-Coda
\>pi:nsi:
\pc:nsi:  *
pensi:  *

/ch[e]que/
\>tsh:ek  *
tshi:k  **
tsh:ek  *

/offs[e]t/
\>sek  *
se:t  *
si:t  **

/B[e]nz/  Id-[high]  Id-Coda  *ET
\>pc:nsi:  *
\pi:nsi:  *
pensi:  *
Phonotactic Constraint-2: Labial

- om
  chloroform  kolofən
  zombie  səŋbej
  ohm  ewmow

- deflect [u] > [ɔ:]

  box  pɔ:kseŋ
  cocktail  kɔ:ktë:
  franc  fələŋə:

  comp  khəmˈlit
  composition  khəmˈphəw
  ping-pong ball  pəŋˈpəm-pə:

- heterorganic NC
  hamburger  hə:nbəw
  rumba  le:nba:

- heterorganic coda filler
  spring  sibətliŋ
  cookie  khəkkəhi:
  Fascism  fətsəisi
  gaberdine  ɡæpˈbaːdiŋ
  romance  ləŋma:nsi
  sauna  sənəː (21)
  mummy  mʊknəi ji