Markedness effects in paradigm reanalysis: Malagasy consonant alternations
Jennifer Kuo, University of California, Los Angeles
jenniferkuo2018@ucla.edu

1 Overview
• How do learners reconstruct a neutralized form?
  - \[ [b\backslash m\backslash a\backslash b] \rightarrow /b\backslash m\backslash a\backslash b/ \]
  - \[ [b\backslash m\backslash a\backslash u] \rightarrow /b\backslash m\backslash a\backslash u/ \]
• Possible factors:
  - distributional information (Ernestus and Baayen, 2003; Albright, 2002)
  - innate biases (Moreton, 2008)
• Paradigm reanalysis as window into phonological learning (Kiparsky, 1965)
• Case study: Malagasy consonant alternations
• Results: effects of markedness bias
  - not predicted by existing models (e.g. Albright, 2002; Nosofsky, 2011)

2 Background: Malagasy
• (C)V syllables, mostly penultimate stress.
• Weak stems: antepenultimate stress (if long enough) and end in “weak syllable” (ka, na, tra [ïaal])
• Weak syllable’s consonant may alternate under suffixation:

<table>
<thead>
<tr>
<th>pattern</th>
<th>stem</th>
<th>passive -(ana)</th>
</tr>
</thead>
<tbody>
<tr>
<td>na</td>
<td>n</td>
<td>a\‘dravina</td>
</tr>
<tr>
<td>m</td>
<td>anâ\‘draina</td>
<td>a\‘dravâna</td>
</tr>
<tr>
<td>ka</td>
<td>h</td>
<td>a\‘gataka</td>
</tr>
<tr>
<td>f</td>
<td>anâha</td>
<td>anahâfana</td>
</tr>
<tr>
<td>t\‘ra</td>
<td>r</td>
<td>i\‘natra</td>
</tr>
<tr>
<td>t</td>
<td>anâ\‘dra tra</td>
<td>ana\‘dratana</td>
</tr>
<tr>
<td>f</td>
<td>a\‘drakutra</td>
<td>a\‘drakufana</td>
</tr>
</tbody>
</table>

• Historically consonant-final (Dahl, 1951; Adelaar, 2012)
  1. Development consonant neutralization
  2. Vowel epenthesis to resolve codas
• Ex: development of tra\‘r alternation:

  *bukID | *bukID-\‘an | Historical
  *wukit  | \(-\rightarrow^{*}\backslash t)\)
  *\‘uwit\‘r  | \(-\rightarrow^{*}\backslash t \rightarrow^{*}\rightarrow^{*}\backslash t)\)
  *\‘uwhit\‘ra | *\‘uwhir\‘an | (Epenthesis)
  \‘uhi\‘tra | \‘uhi\‘rina | Modern

Possible reanalyses for [p\‘akutra]
Direction passive (stem+ana)
\( t \rightarrow r \) pakut-ana\‘pakur-ana
\( r \rightarrow t \) pakur-ana\‘pakut-ana

3 Reanalysis in weak stems
Method: Compare historical and modern Malagasy
• Historical: Austronesian Comparative Dictionary (ACD; Blust and Trussel, 2010)

Predicted vs. Observed reanalyses:
• Distributional models predict reanalysis towards more likely alternant

(1) Historical distribution of alternants (Proto-Malayo-Polynesian).
Reanalysis should be towards historically more frequent alternant (green)

(3) Actual distribution of alternants (modern Malagasy).
Unexpected preference for [r] in modern Malagasy

(4) Documented reanalyses

<table>
<thead>
<tr>
<th>Type</th>
<th>Change</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>na</td>
<td>n→m</td>
<td>3</td>
</tr>
<tr>
<td>ka</td>
<td>h→f</td>
<td>0</td>
</tr>
<tr>
<td>r</td>
<td>t</td>
<td>0</td>
</tr>
<tr>
<td>tra</td>
<td>t→r</td>
<td>33</td>
</tr>
</tbody>
</table>

For tra-final stems, reanalysis is t → r, NOT predicted by distributions

4 Markedness bias
Markedness bias against intervocalic stops explains t → r reanalysis
• Constraint: *\( V-\text{cont}\)V
• Historically, intervocalic lenition in Malagasy (*b>v, *p>f, *d>r, *k, *g>h)
• Typologically common (Kirchner, 1998; Kaplan, 2010; Katz, 2016)
• Active as statistical phonotactic tendency

5 Model Implementation
Goal: Show effect of markedness bias through modeling
Model components:
• MaxEnt Harmonic Grammar (Goldwater and Johnson, 2003) to capture gradient alternations.
• Bias implemented as a Gaussian prior (Wilson, 2006; White, 2013).
• Iterative: Predictions of one iteration is input to next iteration.

Model constraints:
• Constraints enforcing alternation in weak stems: *\( t\text{r} \), *\( k\text{V} \), *\( h\text{V} \)
• Faithfulness constraints: *MAP (Zuraw, 2010, 2013)
• *\( r\text{r} \) enforces r-dissimilation
• *\( \text{V-cont}\)V penalizes tra→r alternation.

Model evaluation:
• Compare models with markedness bias against controls with no bias.
• Results in Section 6

6 Model Results
Result: Reanalysis in Malagasy explained by successive generations of learning modulated by markedness bias

Bias terms: \( \mu \approx \text{preferred weight} \)
• Flat prior (control): uniform \( \mu \)
• P-map (control): For *MAP, perceptually similar mappings get lower \( \mu \)
• Markedness: \( \mu(\backslash V-\text{cont}\text{V}) > \mu(\text{Faith}) \)
• Markedness + P-map

| Candidate | Candidate
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90 (L=6460)</td>
<td>0.98 (L=2233)</td>
</tr>
<tr>
<td>0.68 (L=2793)</td>
<td>0.58 (L=9618)</td>
</tr>
</tbody>
</table>

△ Figure: Models with markedness bias predict more tra→r alternation. (Change in P of tra-final candidates over 10 iterations; PMP: old Malagasy, Mlg: new Malagasy; models whose predictions are closer Mlg are better.)

Takeaway
Markedness effects are found in reanalysis, and can be accounted for by augmenting distributional models with a bias term.

Acknowledgements
Thanks to Bruce Hayes, Kie Zuraw, Claire Moore-Cantwell, members of the UCLA Phonology Seminar for their helpful comments; audience of TripleAFLA for their input; Vololona for her invaluable time and assistance.

References
See next page
References


