Phonological Learning with Output-Driven Maps

Bruce Tesar
Linguistics Dept. / Center for Cognitive Science
Rutgers University, New Brunswick

Eighth North American Phonology Conference
May 9, 2014
Outline

• Simultaneous learning of:
  – constraint rankings
  – underlying forms

• Output-Driven Maps
  – Phono-LOGICAL Reasoning: logical entailment over algebraic lattices

• Exploiting ODM structure in learning
Need Additional Structure

• Joint learning of constraint rankings and underlying forms.
• These techniques are still implausibly slow.
• Faster learning requires additional posited structure:
  – relating the space of possible UFs to rankings.

• Proposal: Output-Driven Maps
A System for Illustration

- **Words:** root + suffix
  - Both roots and suffixes are monosyllabic.
- **Each vowel has two features:**
  - Vowel length: long (+) or short (–)
  - Main stress: stressed (+) or unstressed (–)
- **Example surface words:**
  - Each word has exactly one main stress in the output.
The Constraints

• Six Constraints
  MAINLEFT main stress on the initial syllable
  MAINRIGHT main stress on the final syllable
  *V: no long vowels
  WSP long vowels are stressed
  ID[stress] correspondents have equal stress value
  ID[length] correspondents have equal length value

(McCarthy & Prince 1993, 1995; Prince 1990; Rosenthall 1994)
# Language L20

<table>
<thead>
<tr>
<th>r1=/&gt;pa/</th>
<th>r2=/&gt;pa:/</th>
<th>r3=/&gt;pá/</th>
<th>r4=/&gt;pá:/</th>
<th>s1=/&gt;-ka/</th>
</tr>
</thead>
<tbody>
<tr>
<td>páka</td>
<td>pá:ka</td>
<td>páka</td>
<td>pá:ka</td>
<td>s1=/&gt;-ka/</td>
</tr>
<tr>
<td>páka</td>
<td>pá:ka</td>
<td>páka</td>
<td>pá:ka</td>
<td>s2=/&gt;-ka:</td>
</tr>
<tr>
<td>paká</td>
<td>paká</td>
<td>páka</td>
<td>pá:ka</td>
<td>s3=/&gt;-ká/</td>
</tr>
<tr>
<td>paká:</td>
<td>paká:</td>
<td>páka</td>
<td>pá:ka</td>
<td>s4=/&gt;-ká:/</td>
</tr>
</tbody>
</table>

Ranking: WSP >> ID[s] >> ML >> MR >> ID[l] >> *V:
Output-Driven Maps
(Tesar 2008; Tesar 2014)

• A map is output-driven if:
  – for every grammatical candidate $A \rightarrow X$ of the map:
  – if candidate $B \rightarrow X$ (same output) has greater similarity than $A \rightarrow X$,
  – then $B \rightarrow X$ is also grammatical.

• Simplified:
  – for every grammatical candidate $A \rightarrow X$ of the map:
  – if input $B$ is more similar to $X$ than $A$ is,
  – then $B$ also maps to $X$. 
Greater Similarity

- Candidate B→X has **greater similarity** than candidate A→X if every disparity in B→X has an identical corresponding disparity in A→X.
  - The relation is only defined for pairs of candidates sharing the same output.

(A→X) páká → paká:

\[ [+ - + -] \rightarrow [- - + +] \]

(B→X) paká → paká:

\[ [- - + -] \rightarrow [- - + +] \]
Relative Similarity \((up = \text{greater similarity})\)
Relative Similarity (+/–stress +/–length)
Exploiting ODM Structure in Learning

• ODM structure can be exploited in the learning of both:
  – underlying feature values
  – ranking information

• Major benefit: computational efficiency
Phonotactic Learning

• Identity Map Property
  – underlying forms identical to the observed output.

• The Identity Map property follows from ODM structure.
  – Phonotactic learning can be done as before.
ODM entails the Identity Map Property
Ranking Information Content of *paká*:

<table>
<thead>
<tr>
<th>/paká:/</th>
<th>WSP</th>
<th>ML</th>
<th>MR</th>
<th>*V:</th>
<th>ID[s]</th>
<th>ID[l]</th>
</tr>
</thead>
<tbody>
<tr>
<td>paká:</td>
<td>winner</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>paká</td>
<td>loser</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ERC paká: ~ paká</td>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td>W</td>
<td></td>
</tr>
</tbody>
</table>

Word r1s4 has surface form *paká*:

Mapping that form to itself yields ID[l] ▫ ▫ ▫ *V:
Learning Underlying Feature Values

• ODM: $A \rightarrow X$ entails $B \rightarrow X$
• Contrapositive: NOT ($B \rightarrow X$) entails NOT ($A \rightarrow X$)
  – If a given input cannot map to the output, then all inputs with lesser similarity (additional disparities) cannot map to that output.
Testing Individual Disparities

• Observed output (r1s4): paká:
• What is the underlying length of suffix s4?

• paká→paká: disparity for s4 length only.

• If paká→paká: is inconsistent
  – no other input with s4 set to short maps to paká:
  – s4 can be set to long (+).
Setting s4 to +long
Exponential to Linear

• The learner only needs to test one input for each unset feature.

• Linear in the number of unset features
  – rather than exponential.
Learning Output-Driven Maps

Features Are Set When Contrastive

- *paká→paká*: is inconsistent
- because length is contrastive in stressed position
- which the learner knows via $\text{ID}[l] \gg *V$:
- as determined by phonotactic learning.
Non-Phonotactic Ranking Information

• Find forms in which a set feature is not faithfully realized (Tesar 2006b).
  – Where the feature **alternates**.

• Indicates neutralization.
Unfaithful Features

• Observed output (r3s4): páka

• s4 has already been set to +long.

• Minimal disparity mapping: páka: \(\rightarrow\) páka
Viable Inputs for r3s4
## Ranking Info from r3s4

<table>
<thead>
<tr>
<th>/páka:/</th>
<th>WSP</th>
<th>ML</th>
<th>MR</th>
<th>*V:</th>
<th>ID[s]</th>
<th>ID[l]</th>
</tr>
</thead>
<tbody>
<tr>
<td>páka</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>páka:</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC</td>
<td>W</td>
<td></td>
<td>W</td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td>paká: ~ paká</td>
<td></td>
<td></td>
<td>L</td>
<td>W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>W</td>
<td>L</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

WSP $\gg$ ID[l] $\gg$ *V:

Obtained despite incomplete input knowledge.
Contrast and Neutralization

• Underlying feature values are learned in positions of contrast.

• Non-phonotactic ranking information is learned in positions of neutralization.

• In learning, each feeds the other.
Underlying Forms, Not Surface Allomorphs

<table>
<thead>
<tr>
<th></th>
<th>r1=/pa/</th>
<th>r2=/pa:/</th>
<th>r3=/pá/</th>
<th>r4=/pá:/</th>
<th>s1=/-ka/</th>
<th>s2=/-ka:/</th>
<th>s3=/-ká/</th>
<th>s4=/-ká:/</th>
</tr>
</thead>
<tbody>
<tr>
<td>r1</td>
<td>páka</td>
<td>pá:ka</td>
<td>páka</td>
<td>pá:ka</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r2</td>
<td>páka</td>
<td>pá:ka</td>
<td>páka</td>
<td>pá:ka</td>
<td>s1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r3</td>
<td>paká</td>
<td>paká</td>
<td>páka</td>
<td>pá:ka</td>
<td>s2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r4</td>
<td>paká:</td>
<td>paká:</td>
<td>páka</td>
<td>pá:ka</td>
<td>s3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

r2 always surfaces as pá: or pa (never as pa:)

Bruce Tesar  Linguistics / Center for Cognitive Science  26
## Learning Conspiracies: L9

<table>
<thead>
<tr>
<th></th>
<th>r1=/pa/</th>
<th>r2=/pa:/</th>
<th>r3=/pá/</th>
<th>r4=/pá:/</th>
</tr>
</thead>
<tbody>
<tr>
<td>paká</td>
<td>paká:</td>
<td>paká</td>
<td>paká:</td>
<td>s1=/-ka/</td>
</tr>
<tr>
<td>paká:</td>
<td>paká:</td>
<td>paká:</td>
<td>paká:</td>
<td>s2=/-ka:/</td>
</tr>
<tr>
<td>paká</td>
<td>paká:</td>
<td>paká</td>
<td>paká</td>
<td>s3=/-ká/</td>
</tr>
<tr>
<td>paká:</td>
<td>paká:</td>
<td>paká:</td>
<td>paká:</td>
<td>s4=/-ká:/</td>
</tr>
</tbody>
</table>

- **r1s1**: /paka/ → paká  
  Default final stress

- **r2s3**: /pa:ká/ → pá:ka  
  WSP, via stress shift

- **r4s2**: /pá:ka:/ → pá:ka  
  WSP, via vowel shortening
L9 Phonotactic Learning

• L9 includes contrasts in stress and length.

• ID[l] $\gg$ *V: (contrast in length)
  – paká paká:

• ID[s] $\gg$ {ML,MR} (contrast in stress)
  – paká páka
L9 UF Learning(1)

• r1s2: /paká/ cannot map to paká:
  – ID[l] $\gg *V$:

• s2 can be set underlyingly to +long.
  – Because –long is inconsistent.
  – s2 now has lexical entry /?,+/
L9 Non-phonotactic Ranking Info(1)

- r4s2 surfaces as \( \text{pá:ka} \)
  - s2 surfaces as –long
  - \(/\text{pá:ka}/ \rightarrow \text{pá:ka}\)

- \( \text{WSP} \gg \text{ID}[l] \) (vowel shortening)
L9 Contrast Pair UF Learning

• r1s1 paká
• r3s1 páka

• r1 and r3 must contrast in underlying stress.
  – Set r1 to –stress.
  – Set r3 to +stress.
L9 Non-phonotactic Ranking Info(2)

- r3s3 surfaces as paká
  - r3 surfaces as –stress.
  - /páká/ → paká

- MR ≫ ML (default final stress)
L9 UF Learning(2)

• r4s2: /pá:ká:/ cannot map to pá:ka
  – MR ⪰ ML

• s2 can be set underlyingly to –stress.
  – Because +stress is inconsistent.
  – s2 now has lexical entry /−,+/

Bruce Tesar  Linguistics / Center for Cognitive Science
L9 Non-phonotactic Ranking Info (3)

- r3s2 surfaces as *paká*:
  - s2 surfaces as +stress.
  - /páka:/ → paká:

- ID[l] ≫ ID[s]  (stress shift)
Just Another Grammar

• No special mechanisms for learning conspiracies.

• No special mechanisms for non-allomorphic UF's.
Conclusions

• ODM structure makes much more efficient learning possible.
  – Reduction from exponential to linear.
  – Both underlying forms and ranking information.

• Phono-LOGICAL reasoning: entailment over algebraic lattices.

• Jointly leveraging the two forms of paradigmatic information.
  – contrast
  – alternation