Role of perceptual similarity in learning phonological alternations

James White
University of Ottawa
Role of perceptual similarity in phonological learning?

• Does perceptual similarity play a role during phonological learning?
  – If so, what is that role?

• Part of a larger question: Does phonetic substance play a role in phonological learning?
  – Are there substantive biases\(^1\)?

1. Wilson, 2006
Steriade’s P-map proposal

- **P-map**: a mental representation of the relative perceptual similarity between speech sounds in a given context.¹
  - Perhaps based on an individual’s prior perceptual experience.

- Steriade proposed that learners are biased by the P-map when learning phonological patterns.
  - Phonological processes assumed, *a priori*, to involve minimal perceptual modification.

1. Steriade, 2001/2008
Today’s talk

• **Goal**: Investigate this issue by comparing learning models both with and without a similarity bias.

• **Test case**: Experimental data showing that adult learners disprefer *saltatory alternations* (White, 2014, *Cognition*).

  = a type of alternation involving not minimal change, but excessive change.
Roadmap

1. Overview of saltatory alternations

2. Overview of experimental results

3. Modeling
1. Saltatory alternation
Saltatory alternation

• Phonological alternation where an intermediate sound is “leaped over”.¹

• Campidanian Sardinian:²
  – /p/ → [β] / V ___V  
  – no change for /b/.
  
  [päi] → [s:u βäi]  ‘the bread’
  [bĩu] → [s:u bĩu]  ‘the wine’

1. White, 2014
2. Bolognesi, 1998
Saltatory alternation

- Phonological alternation where an intermediate sound is “leaped over”.\(^1\)
- Campidanian Sardinian:\(^2\)

---

1. White, 2014
2. Bolognesi, 1998
Saltatory alternation

- Phonological alternation where an intermediate sound is “leaped over”.¹
- Campidanian Sardinian:²

1. White, 2014
2. Bolognesi, 1998
2. Experimental results

Learners prefer to avoid saltatory alternations.\textsuperscript{1}

\textsuperscript{1} White, 2014
Artificial language experiments (adult learners)

1. Exposure phase

[kamap]  [kamavi]

Also: non-changing fillers like [luman]  [lumani]
Artificial language experiments (adult learners)

1. Exposure phase
   - [kamap]

2. Verification phase
   - [kamap]

3. Generalization phase
   - [lunub]
   or
   - [kamapi]
   or
   - [kamavi]??

   - [lunubi]
   or
   - [lunuvi]???
Crucial results to be accounted for

Ambiguous Saltation

Test condition

Control condition

Explicit Saltation
3. Maximum entropy learning model

For previous uses, see Goldwater & Johnson, 2003; Wilson, 2006; Hayes & Wilson, 2008; Hayes et al., 2009; Martin, 2012; others. Implemented using the MaxEnt Grammar Tool (available at Bruce Hayes’s webpage).
2 things the learning model should account for

1. Saltatory alternations exist and thus must be learnable for the child.

2. This type of alternation is dispreferred by learners.
   
   – Can we model the experimental results?
Overview of the model

- Set of OT-style constraints
- Input forms (singular words)
- Candidate output forms (plural words)
- Constraint violations

Provide the model

Training data
(observed input-output pairs)

Prior
(a priori preferred weights)

Grammar
(=weighted constraints)

Predicted probability
of each output

(Maxent learning)
• Feature-based markedness constraints, which motivate alternating.
  – *V [–voice] V
  – *V [–continuant] V

• Correspondence constraints banning alternations between specific pairs of sounds.¹
  – E.g., *MAP(p,v) = don’t have an alternation between [p] and [v].

Why *MAP constraints?

• Traditional faithfulness constraints (classical OT) cannot generate saltation.\(^1,2\)
  – /p/ → [v], when [b] is legal, results in a gratuitous faithfulness violation.
  – True, even with weighted constraints.

• Straightforward implementation of the similarity bias.

## Training data

- Same as the training data in the experiments.

- Ambiguous Saltation experiment:

<table>
<thead>
<tr>
<th>input</th>
<th>possible outputs</th>
<th># observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>v</td>
<td>18</td>
</tr>
<tr>
<td>p</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>f</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>input</th>
<th>possible outputs</th>
<th># observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>δ</td>
<td>18</td>
</tr>
<tr>
<td>t</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>θ</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>d</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Prior (= the bias)

• Biases the learning towards certain outcomes, based on *a priori* assumptions.
  – In this case, based on perceptual similarity.
  – *Soft bias*, not an absolute restriction.

• The prior is Gaussian. Provide:
  – $\mu = \text{preferred weight for each constraint}$.
  – $\sigma^2 = \text{how tightly constraints are held to their preferred weight}$. 
Measure of similarity

- How people judge the similarity of speech sounds is likely complex.\(^1, 2, 3\)
- One possibility:

\[ \begin{align*}
\text{Featural similarity} \\
\text{voicing} &  \quad \longleftrightarrow \\
p & \quad \downarrow 1 \quad b \\
\downarrow \text{continuancy} & \quad 1 \\
f & \quad \longleftrightarrow 1 \\
\downarrow \text{continuancy} & \quad \longleftrightarrow \\
v & \quad 1 \\
\end{align*} \]

1. Steriade, 2001/2008  
3. Cristia et al., 2013
Perceptual similarity

• I use **mutual confusability** as a simplified measure of perceptual similarity.
  
  – Data from published confusion experiments with adult English speakers.¹

• Example:

1. Wang & Bilger, 1973
• **Goal**: go directly from confusion proportions to prior constraint weights. (*No cherry-picking weights*).

• **Solution**: Train up a separate maxent model intended solely to generate prior weights \((\mu)\), based on confusion probabilities.

• **Intuitively**: Represents the listener’s perceptual experience, a computational version of the P-map.
## Prior weights

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Substantively Biased</th>
</tr>
</thead>
<tbody>
<tr>
<td>*V [-voice] V</td>
<td>0</td>
</tr>
<tr>
<td>*V [-cont] V</td>
<td>0</td>
</tr>
<tr>
<td>*MAP( p, f )</td>
<td>1.34</td>
</tr>
<tr>
<td>( p, b )</td>
<td>2.44</td>
</tr>
<tr>
<td>( p, v )</td>
<td>3.65</td>
</tr>
<tr>
<td>( b, v )</td>
<td>1.30</td>
</tr>
<tr>
<td>( b, f )</td>
<td>1.96</td>
</tr>
<tr>
<td>( f, v )</td>
<td>2.56</td>
</tr>
</tbody>
</table>
2 other models compared

1. **Anti-alternation**: All *MAP constraints start with the same weight.
   - prior weight = **average** of prior weights in the substantively biased model.
   - *A priori* preference to avoid any alternation equally, regardless of similarity.

2. **Unbiased**: Prior weight of 0 for all constraints.
## Prior weights

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Substantively Biased</th>
<th>Anti-alternation</th>
<th>Unbiased</th>
</tr>
</thead>
<tbody>
<tr>
<td>*V [-voice] V</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>*V [-cont] V</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>*MAP( p, f )</td>
<td>1.34</td>
<td>2.27</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2.44</td>
<td>2.27</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3.65</td>
<td>2.27</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1.30</td>
<td>2.27</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1.96</td>
<td>2.27</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2.56</td>
<td>2.27</td>
<td>0</td>
</tr>
</tbody>
</table>
Learning procedure

• Find the set of constraint weights that maximizes this function:

\[
\left[ \sum_{j=1}^{n} \log \Pr(y_j \mid x_j) \right] - \left[ \sum_{i=1}^{m} \frac{(w_i - \mu_i)^2}{2\sigma_i^2} \right]
\]

Maximize the likelihood of the training data with a penalty for weights that vary from the prior

• Model will provably always succeed at finding the “best” grammar by this criterion.\(^1\)

1. Berger et al., 1996
Returning to the crucial experimental results

Ambiguous Saltation

Test condition

- p to f: 45% (P: 41%)
- f to v: 64% (P: 70%)

Control condition

- p to b: 21% (P: 22%)
- b to p: 7% (P: 12%)

Explicit Saltation

- p to b: 21% (P: 27%)
- b to p: 7% (P: 7%)
Model performance

Substantively biased model

Recall: No manipulation of prior weights by hand!
Compared to other models

Substantively biased: $r^2 = .95$

Anti-alternation: $r^2 = .67$

Unbiased: $r^2 = .25$
Where do the control models go wrong?

• Unbiased model:
  – In general, predicts too much generalization – no reason to avoid positing new alternations.

• Anti-alternation model:
  – Unable to account for subtler variations between segments (e.g., [b] changed to [v] more often than [f]).
  – Ends up predicting more generalization in the control case than in the saltation case – opposite of the actual results!!
Discussion

• Anti-alternation model outperforms Unbiased model.
  – Consistent with a general preference for avoiding alternations (e.g., OO-Faith set high by default).\(^1,\(^2\)

• Substantively biased model outperforms Anti-alternation model.
  – Evidence that perceptual similarity plays a role even beyond a general preference to avoid alternations.

• General framework for looking at the role of perceptual similarity in phonological learning.

Prior as a soft bias

• The prior is crucial to the model’s success. It allows the desired learning pattern:
  – Alternations between dissimilar sounds are initially dispreferred.
  – But with enough training data, the prior can be overcome – they are learnable.

• The anti-saltation effect seen in the experiments seems to fall out from a more general similarity bias.
Evidence with infant learners?

• 12-month-olds learning potentially saltatory alternation generalize to intermediate sounds.$^{1}$

• 12-month-old English-learning infants know $[d \sim r]$, but not $[t \sim r]$, despite greater support for $[t \sim r]$ in their input.$^{2}$

2. Sundara et al., 2013 (BUCLD)
Thank you!

• Acknowledgments:
  – For help and discussion, special thanks to Bruce Hayes, Megha Sundara, Kie Zuraw, Robert Daland, Sharon Peperkamp, Adam Albright, and audiences at UCLA, the University of Ottawa, and Stony Brook University.
  – Thanks to my undergraduate research assistants as well as research assistants at the UCLA Language Acquisition Lab.
Creating the prior (details)

• **Input** = identification data from confusion experiments.¹

  E.g.,

  \[
  \begin{array}{cccc}
  \text{Stimulus} & \text{p} & \text{b} & \text{f} & \text{v} \\
  \hline
  \text{p} & 1844 & 54 & 159 & 26 \\
  \text{b} & 206 & 1331 & 241 & 408 \\
  \text{f} & 601 & 161 & 1202 & 93 \\
  \text{v} & 51 & 386 & 127 & 1428 \\
  \end{array}
  \]

  \[
  \begin{array}{cccc}
  \text{Stimulus} & \text{t} & \text{d} & \theta & \delta \\
  \hline
  \text{t} & 1765 & 107 & 92 & 26 \\
  \text{d} & 91 & 1640 & 75 & 193 \\
  \theta & 267 & 118 & 712 & 135 \\
  \delta & 44 & 371 & 125 & 680 \\
  \end{array}
  \]

• Each *MAP* constraint is weighted according to how likely its sounds are to be confused for each other.

• **Output weights** → **Prior of main model**

¹. Wang & Bilger 1973
# Effect of different confusion data

<table>
<thead>
<tr>
<th>Source</th>
<th>Table #</th>
<th>Context</th>
<th>In noise?</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WB 1973</td>
<td>2-3</td>
<td>CV and VC</td>
<td>white noise</td>
<td>.95</td>
</tr>
<tr>
<td>WB 1973</td>
<td>2</td>
<td>CV</td>
<td>white noise</td>
<td>.93</td>
</tr>
<tr>
<td>WB 1973</td>
<td>3</td>
<td>VC</td>
<td>white noise</td>
<td>.92</td>
</tr>
<tr>
<td>WB 1973</td>
<td>6-7</td>
<td>CV and VC</td>
<td>none</td>
<td>.93</td>
</tr>
<tr>
<td>WB 1973</td>
<td>6</td>
<td>CV</td>
<td>none</td>
<td>.82</td>
</tr>
<tr>
<td>WB 1973</td>
<td>7</td>
<td>VC</td>
<td>none</td>
<td>.96</td>
</tr>
<tr>
<td>MN 1955</td>
<td>2-6</td>
<td>CV</td>
<td>white noise</td>
<td>.94</td>
</tr>
<tr>
<td>C-etal 2004</td>
<td>----</td>
<td>CV and VC</td>
<td>babbled noise</td>
<td>.82</td>
</tr>
<tr>
<td>C-etal 2004</td>
<td>----</td>
<td>CV</td>
<td>babbled noise</td>
<td>.79</td>
</tr>
<tr>
<td>C-etal 2004</td>
<td>----</td>
<td>VC</td>
<td>babbled noise</td>
<td>.77</td>
</tr>
<tr>
<td><strong>Unbiased model (for comparison)</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>.25</strong></td>
</tr>
</tbody>
</table>
How the weights change during learning

### Explicit Saltation experiment

\( p \rightarrow v, b \rightarrow b \)

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Prior weight</th>
<th>Post-learning weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>*V [-voice] V</td>
<td>0</td>
<td>2.45</td>
</tr>
<tr>
<td>*V [-cont] V</td>
<td>0</td>
<td>1.05</td>
</tr>
<tr>
<td>*MAP( p, f )</td>
<td>1.34</td>
<td>1.74</td>
</tr>
<tr>
<td>( p, b )</td>
<td>2.44</td>
<td>2.94</td>
</tr>
<tr>
<td>( p, v )</td>
<td>3.65</td>
<td>1.96</td>
</tr>
<tr>
<td>( b, v )</td>
<td>1.30</td>
<td>2.02</td>
</tr>
<tr>
<td>( b, f )</td>
<td>1.96</td>
<td>2.02</td>
</tr>
<tr>
<td>( f, v )</td>
<td>2.56</td>
<td>2.56</td>
</tr>
</tbody>
</table>
Effect of $\sigma$

Prior has almost no practical effect on the constraint weights, leaving the weights to be almost entirely dictated by the training data. As a result, all three models converge at (virtually) the same predictions and thus have very similar $r^2$ values.

Figure 11. Proportion variance explained ($r^2$) by the substantively biased model, the high faith model, and the unbiased model, according to the value of $\sigma^2$.  

4.4.5 Effect of different confusion matrices used to derive the prior

Given that the prior weights are calculated directly from confusion probabilities, the prior weights, and thus the model's performance, will vary depending on which confusion data are used as input. Implicit in this design is the prediction that real language learners will learn slightly differently depending on their own perceptual experience and language background. This strikes me as a reasonable assumption.
Figure 12. Predictions of the substantively biased model plotted against the experimental results from the production experiment. Overall $r^2 = .87$. 

For comparison, the predictions of the unbiased model are plotted against the experimental results in Figure 13. For the best comparison, the unbiased model was also augmented with an *A LTERNATE constraint with the same prior values that it had in the biased model (i.e., $\mu$ of 1 and $\sigma$ of .0001). As the plot shows, the unbiased model is much less successful at predicting the experimental results; as a result, the fit between the model predictions and the experimental results is much lower ($r^2 = .39$).
Predicting probabilities from the grammar

Explicit Saltation experiment

<table>
<thead>
<tr>
<th>/kama(p/</th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>kama(v/</td>
<td>2.45</td>
<td>2.02</td>
<td>1.96</td>
<td>1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kamai</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kama(p/</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Predicting probabilities from the grammar

**Explicit Saltation experiment**

<table>
<thead>
<tr>
<th>/kamap/</th>
<th>*V[-voice]V 2.45</th>
<th>*MAP(b, v) 2.02</th>
<th>*MAP(p, v) 1.96</th>
<th>*V[-cont]V 1.05</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>kamavi</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kamapi</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Input form and output candidates**
**Predicting probabilities from the grammar**

**Explicit Saltation experiment**

<table>
<thead>
<tr>
<th>Constraint weights</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>/kama</em></td>
<td>2.45</td>
<td></td>
</tr>
<tr>
<td><em>/kama</em></td>
<td>2.02</td>
<td></td>
</tr>
<tr>
<td><em>/kama</em></td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td><em>/kama</em></td>
<td>1.05</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/kama/</th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>kamavi</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kamapi</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

**Training:**

- $p$
- $b$
- $v$
**Predicting probabilities from the grammar**

Explicit Saltation experiment

<table>
<thead>
<tr>
<th>/kamap/</th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>kamavi</td>
<td>2.45</td>
<td></td>
<td>1.96</td>
<td></td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>kamaπi</td>
<td>2.45</td>
<td>1.96</td>
<td></td>
<td>1.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Training:

- **p**
- **b**
- **v**
Predicting probabilities from the grammar

Explicit Saltation experiment

<table>
<thead>
<tr>
<th>/kama(p/)</th>
<th>*V[-voice]V 2.45</th>
<th>*MAP(b, v) 2.02</th>
<th>*MAP(p, v) 1.96</th>
<th>*V[-cont]V 1.05</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>kamavi</td>
<td></td>
<td></td>
<td>1.96</td>
<td></td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>kamapi</td>
<td>2.45</td>
<td></td>
<td></td>
<td>1.05</td>
<td>3.50</td>
<td></td>
</tr>
</tbody>
</table>

Training: \(p \Rightarrow b \Rightarrow v\)

\(\sum e (\neg \text{penalty})\)

**Sum**
## Predicting probabilities from the grammar

### Explicit Saltation experiment

<table>
<thead>
<tr>
<th>/kamaₚ/</th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>kamavi</td>
<td>2.45</td>
<td></td>
<td>1.96</td>
<td>1.96</td>
<td></td>
<td>82 %</td>
</tr>
<tr>
<td>kamaₚi</td>
<td>2.45</td>
<td></td>
<td>1.05</td>
<td>3.50</td>
<td></td>
<td>18 %</td>
</tr>
</tbody>
</table>

Training: p → b → v
## Predicting probabilities from the grammar

### Explicit Saltation experiment

<table>
<thead>
<tr>
<th></th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>/kama</strong>/</td>
<td>2.45</td>
<td>2.02</td>
<td>1.96</td>
<td>1.05</td>
<td>3.50</td>
<td><strong>18 %</strong></td>
</tr>
<tr>
<td>kamavi</td>
<td></td>
<td></td>
<td>1.96</td>
<td></td>
<td>1.96</td>
<td><strong>82 %</strong></td>
</tr>
<tr>
<td>kamapi</td>
<td>2.45</td>
<td></td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Explicit Saltation experiment

<table>
<thead>
<tr>
<th></th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>/lunu</strong>/</td>
<td>2.45</td>
<td>2.02</td>
<td>1.96</td>
<td>1.05</td>
<td>2.02</td>
<td><strong>27 %</strong></td>
</tr>
<tr>
<td>lunuvi</td>
<td></td>
<td></td>
<td>2.02</td>
<td></td>
<td>2.02</td>
<td></td>
</tr>
<tr>
<td>lunubi</td>
<td></td>
<td></td>
<td>1.05</td>
<td></td>
<td>1.05</td>
<td><strong>73 %</strong></td>
</tr>
</tbody>
</table>
Predicting probabilities from the grammar

Explicit Saltation experiment

<table>
<thead>
<tr>
<th></th>
<th>*V[-voice]V</th>
<th>*MAP(b, v)</th>
<th>*MAP(p, v)</th>
<th>*V[-cont]V</th>
<th>Total penalty</th>
<th>Predicted output</th>
</tr>
</thead>
<tbody>
<tr>
<td>/kamap/</td>
<td>2.45</td>
<td>2.02</td>
<td>1.96</td>
<td>1.05</td>
<td>1.96</td>
<td>82 %</td>
</tr>
<tr>
<td>kamavi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kamapi</td>
<td>2.45</td>
<td></td>
<td>1.05</td>
<td></td>
<td>3.50</td>
<td>18 %</td>
</tr>
<tr>
<td>/lunub/</td>
<td>2.45</td>
<td>2.02</td>
<td>1.96</td>
<td>1.05</td>
<td>2.02</td>
<td>27 %</td>
</tr>
<tr>
<td>lunuvi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lunubi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.05</td>
<td>73 %</td>
</tr>
</tbody>
</table>

Training: $p \xrightarrow{b} v$
Exp. 1 Results

Potential Saltatory

Control

Percent changing plural option chosen

0 10 20 30 40 50 60 70 80 90 100

Stops  Fricatives  Stops  Fricatives
Exp. 2 Results

Saltatory | Control

Percent changing plural option chosen (errors)

0  10  20  30  40  50  60  70  80  90  100

Stops sub-group | Fricatives sub-group
Stops sub-group | Fricatives sub-group