Autologistic Regression in Linguistic Typology
Yoshihiko Asao
yoshihiko@buffalo.edu

Introduction
- A typological frequency difference is often taken as a linguistic preference and given linguistic explanations.
- However, there are often large-scale geographical patterns.
- It is difficult to distinguish a true linguistic preference from a historical accident.

Previous Approaches
- Independent sample approach (Perkins 1989, among others)
  - Needs to discard most of the data.
  - Arbitrariness and potential interdependence between language areas.
  - Finer-grained geographical information is discarded.

Autologistic Regression
- Similar to the logistic regression in Bickel (2008).
- Instead of language areas, the opinions from neighbors are a part of the model.
- Inspired by discussions on similar issues in ecology (Dormann 2007).

Autologistic Regression (cont.)
- Key idea: If most variance is explained away by the retention of the features of related languages, the evidence for the universal linguistic preference is weak.
- More predictable from neighbors: The pattern can be explained by a few historical accident.
- Less predictable from neighbors: Spontaneous change is frequent, suggesting the frequency difference is more likely to be a linguistic preference.

Procedure
- Data values and geographical distances are taken from WALS chapters (Dryer and Haspelmath 2011).
- Find the best model using stepwise regressions with AIC.

Results: Implicational Universals

<table>
<thead>
<tr>
<th>NRel if VO</th>
<th>NRel</th>
<th>RelN</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO</td>
<td>416</td>
<td>5</td>
</tr>
<tr>
<td>OV</td>
<td>113</td>
<td>132</td>
</tr>
</tbody>
</table>

Correlation of NRel and VO

N/Rel = α + β₁V/O + β₂N + β₃(V/O*N) + ϵ

TH-SOUND (non-sibilant dental or alveolar fricative)

<table>
<thead>
<tr>
<th>AIC</th>
<th>pR²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Th-sound ~ I</td>
<td>308.5</td>
</tr>
<tr>
<td>Th-sound ~ Neighbor</td>
<td>303.5</td>
</tr>
</tbody>
</table>

Effect of neighbors

Effect of basic word order (Correlation of NRel and VO)

World average

Conclusion
Autologistic regression may be a useful method to discern a true linguistic preference from a historical accident.

References
- Dryer (1992) Greenbergian word order correlations.