Explorations in creole research with phylogenetic tools

Aymeric Daval-Markussen & Peter Bakker
Research Centre for Grammar and Language Use,
University of Aarhus
Phylogenetic methods in linguistics

• Gray & Atkinson 2003; McMahon & McMahon 2003;
• Dunn et al. 2005, 2008; Gray et al. 2009...
  and many more since!
• Recently applied to creoles Cysouw 2009; Bakker et al. 2011 ; Daval-Markussen 2011
Why study creoles phylogenetically?

• Not a phylum in the traditional sense
• Evolutionary path characterized by break in transmission
• Creolization processes provide a unique window on human linguistic capacities
Applying phylogenetic tools to creoles: a case-study

- English-based creoles: largest and best-studied group
- Structural as well as lexical features
- Phylogenetic tools to visualize linguistic relationships
- Neighbor-Net networks rather than trees
33 English-based creoles, 122 structural and lexical features
The challenges of creolistics

- Superstrates/lexifiers
- Substrates
  - Relexification hypothesis (e.g. Lefebvre)
  - Berbice Dutch and Ijo
- Feature pool hypothesis
- Restructuring/SLA universals
Some testable hypotheses

• Superstrates: cluster with respective creoles
• Substrates: cluster with the creoles whose formation they have influenced
• Feature Pool: all languages will be drawn towards the creole in which formation they were involved
• Universals: no typological patterning except for creoles
Comparative Creole Syntax
Holm & Patrick 2007

• 97 morphosyntactic features
• 18 creoles worldwide
• 7 lexifiers (2 non-Indo-European)

• The most extensive database available to date
Superstrates: 18 creoles and 7 lexifiers
Substrates: 18 creoles and 19 substrates
The Feature Pool: 18 creoles, 3 substrates and 1 lexifier
Universals: 18 creoles, 19 substrates, 7 lexifiers, 8 non-creoles
Creole typology

- Creoles as a typological group
- ‘Creole Exceptionalism’
- Neighbor-Joining trees rather than networks
- Structural features only
- WALS languages and features only
50 languages, 18 binary features
Moving away from CCS

- Getting rid of redundant and interdependent features
- WALS database richer
- Finer-grained distinctions with multi-state characters
- Multi-state characters less likely to reflect chance distribution than binary encoding
- 100% cells filled
Matching CCS features with WALS features: 9 features shared by a majority of CCS creoles

<table>
<thead>
<tr>
<th>Description</th>
<th>CCS feature</th>
<th>WALS feature</th>
<th>Shared % CCS creoles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indefinite article</td>
<td>15.2</td>
<td>38</td>
<td>100%</td>
</tr>
<tr>
<td>TMA markers</td>
<td>2</td>
<td>69</td>
<td>100%</td>
</tr>
<tr>
<td>Negation</td>
<td>10.1</td>
<td>112</td>
<td>88.89%</td>
</tr>
<tr>
<td>Predicative possession</td>
<td>13.6</td>
<td>117</td>
<td>88.89%</td>
</tr>
<tr>
<td>Order of Adj-N</td>
<td>15.9</td>
<td>87</td>
<td>66.67%</td>
</tr>
<tr>
<td>Comparison</td>
<td>12.7</td>
<td>121</td>
<td>66.67%</td>
</tr>
<tr>
<td>NP conjunction</td>
<td>18.1</td>
<td>63</td>
<td>61.11%</td>
</tr>
<tr>
<td>Order of Gen-N</td>
<td>16.2</td>
<td>86</td>
<td>61.11%</td>
</tr>
<tr>
<td>Passive</td>
<td>11.2</td>
<td>107</td>
<td>61.11%</td>
</tr>
</tbody>
</table>
61 languages, 9 multi-state features
76 languages, 6 multi-state features
134 languages, 4 multi-state features
Next steps

- Identify a third feature (in a similar fashion)
- Include more non-creoles: WALS updates, other available databases
- Include more creoles: APiCS (available late 2012)
- Include a diachronic dimension
- Apply other methods to visualize evolutionary processes proper to creolization
Thanks for your attention!